VERSITA

Mathematica Slovaca

DOI: 10.2478/s12175-008-0081-4
Math. Slovaca **58** (2008), No. 3, 377–386

ON GENERALIZED MEASURES OF ENTROPY AND DEPENDENCE

Parmil Kumar* — D. S. Hooda**

(Communicated by Anatolij Dvurečenskij)

ABSTRACT. Various authors have studied extensions of Shannon's entropy but their inferential properties and applications in applied sciences have not invited proper attention from researchers. In the present paper we explore the motivation and implication of using various classes of the generalized entropies and conditional entropies. We evaluate β-class and (α,β)-class entropies for multivariate normal density function. We also obtain the measures of dependence in terms of the classes of generalized entropies.

1. Introduction

Since 1948 many extensions of Shannon's entropy have been studied by various authors. Some of these have been thoroughly discussed by [3]. However, most of these extensions are purely mathematical formulations and not much is known about their inter-relationships, inferential properties and their applications in other areas like statistics, thermo-dynamics, accountancy, image processing, reliability, finance, economics and pattern recognition.

It has been observed that most of the researchers have focussed mainly on Shannon's entropy, perhaps because of its simplicity. However, we explore the motivation and implication of using various generalized classes of entropies and conditional entropies due to [2], [4] and [6].

The use of above mentioned generalized entropies may lead to statistical results different to those obtained by using Shannon's entropy. Ullah [7] has given some general results related to β-class entropies. Gokhale [1] discussed the joint and conditional entropies and has shown that if conditional densities

2000 Mathematics Subject Classification: Primary 94A15, 94A17.
Keywords: Shannon's entropy, conditional entropy, measure of dependence, covariance matrix, exponential distribution.

of a bivariate random variable have maximum entropies subject to certain constraints, then the bivariate density also maximizes entropy subject to appropriate constraints.

In Section 2, we discuss generalized measures of entropies and conditional entropies. In Section 3, we evaluate β-class and (α, β)-class entropies for multivariate normal density function. In Section 4 we obtain the measures of dependence in terms of these generalized classes of entropies.

2. Generalized classes of entropies and conditional entropies

Let us consider a random vector X $= (x_1, x_2, \dots, x_n)$ with the multivariate probability density function $f(x)$ such that $\int\limits_{x \in X} f(x)\,\mathrm{d}x = 1$. Then information content in observations taken from the population having the probability density function $f(x)$ is given by $-\log f(x)$ and the expected information in X as given by [5] is

$$H(f) = -E\big[\log f(x)\big] = -\int\limits_X f(x) \log f(x)\,\mathrm{d}x. \tag{2.1}$$

Thus, $H(f)$ in (2.1) is a measure of average information or uncertainty removed. The use of the notion $H(f)$ emphasizes that the entropy depends on the probability density function of X rather than its actual values.

As described earlier, Shannon's entropy is the expected value of the function $\phi(f) = -\log f$, which satisfies $\phi(1) = 0$ and $\phi(0) = \infty$. In general, one can choose any convex function with the condition that $\phi(1) = 0$, as a measure of information. So expected information content is given by

$$H(f) = E\big[\phi(f)\big] = \int\limits_X \phi(f) f\,\mathrm{d}x. \tag{2.2}$$

These are called a class of ϕ-entropies (refer to [7]). Axiomatic approach has been used in the literature to choose the information function $\phi(\cdot)$ and some of them are discussed here. Ullah [7] considered a β-class of function given by

$$\phi_\beta(f) = \begin{cases} \frac{1}{\beta-1}(1 - f^{\beta-1}), & \beta \neq 1,\ \beta > 0, \\ -\log f, & \beta = 1 \end{cases} \tag{2.3}$$

where β is a non-stochastic constant. Thus, we write a class of generalized entropy measure as

$$H_\beta(f) = \begin{cases} \frac{1}{\beta-1}\left[1-\int f^\beta \,\mathrm{d}x\right] \\ \frac{1}{\beta-1}\left[1-E(f^{\beta-1})\right], & \beta \neq 1,\ \beta > 0, \\ -E\log f, & \beta = 1. \end{cases} \tag{2.4}$$

This is called the β-class entropy which was initially proposed and studied by [2]. It may be seen that in case $\beta \to 1$, (2.4) reduces to (2.1).

Next we consider another (α,β)-class of entropy measures, which is a generalization of (2.4) and is given by

$$\begin{gathered} H_{\alpha,\beta}(f) = \frac{1}{\beta-\alpha}\left[\int_X f^\alpha \,\mathrm{d}x - \int_X f^\beta \,\mathrm{d}x\right], \\ \alpha \neq \beta, \quad 0 < \alpha \leq 1,\ \beta > 1 \text{ or } 0 < \beta \leq 1,\ \alpha > 1. \end{gathered} \tag{2.5}$$

Consider a bivariate case $X = [x_1, x_2]$ and let $f(x)$ be the joint density function, and $f_1(x_1)$ and $f_2(x_2)$ be their marginal densities, then

$$\begin{aligned} H(f) &= -\iint f(x_1,x_2)\log f(x_1,x_2)\,\mathrm{d}x_1\mathrm{d}x_2 \\ &= -\int f(x_1)\log f(x_1)\,\mathrm{d}x_1 - \int f(x_2/x_1)\log f(x_2/x_1)\,\mathrm{d}x_2 \\ &= H(f_1) + H(f_2/f_1), \end{aligned} \tag{2.6}$$

where $H(f_1)$ is the entropy of f_1 and $H(f_2/f_1)$ is the conditional entropy of the density function f_2 given f_1.

It may be noted that $H(f_1/f_2) \neq H(f_2/f_1)$. However,

$$H(f_1) + H(f_2/f_1) = H(f) = H(f_2) + H(f_1/f_2)$$

or

$$H(f_1) - H(f_1/f_2) = H(f_2) - H(f_2/f_1). \tag{2.7}$$

It implies that

$$H(f_2) - H(f_2/f_1) \geq 0. \tag{2.8}$$

The sign of equality holds if x_1 and x_2 are independent. Thus (2.8) can be considered as a measure of dependence.

In case $X = [x_1, x_2, \ldots, x_n]$, we get

$$H(f) = \sum_{i=1}^{n} H(f_i/f_{i-1},\ldots,f_1) \leq \sum_{i=1}^{n} H(f_i), \tag{2.9}$$

where equality holds iff $x_1, x_2, \ldots, x_n$ are independent.

Next, we study conditional entropies of the following β-class function:

$$\begin{aligned}
H_\beta(f) &= \frac{1}{\beta-1}\left[1 - \int\int f^\beta(x_1,x_2)\,\mathrm{d}x_1\mathrm{d}x_2\right] \\
&= \frac{1}{\beta-1}\left[1 - \int_{x_1} f^\beta(x_1)\left\{\int_{x_2} f^\beta(x_2/x_1)\,\mathrm{d}x_2\right\}\mathrm{d}x_1\right] \\
&= \frac{1}{\beta-1}\left[1 - \int_{x_1} f^\beta(x_1)\,\mathrm{d}x_1 + \int_{x_1} f^\beta(x_1)\,\mathrm{d}x_1\right. \\
&\qquad\qquad \left. - \int_{x_1} f^\beta(x_1)\left\{\int_{x_2} f^\beta(x_2|x_1)\,\mathrm{d}x_2\right\}\mathrm{d}x_1\right] \\
&= H_\beta(f_1) + \int_{x_1} f^\beta(x_1)\left[1 - \int_{x_2} f^\beta(x_2|x_1)\,\mathrm{d}x_2\right]\mathrm{d}x_1 \\
&= H_\beta(f_1) + \int_{x_1} f^\beta(x_1) H_\beta(f_2|f_1)\,\mathrm{d}x_1 \\
&= H_\beta(f_1) + E_{f_1}\left[f_1^{\beta-1} H_\beta(f_2|f_1)\right].
\end{aligned} \tag{2.10}$$

In case (α,β)-class entropy we have

$$\begin{aligned}
H_{\alpha,\beta}(f) &= \frac{1}{\beta-\alpha}\left[\int_{x_1}\int_{x_2} f^\alpha(x_1,x_2)\,\mathrm{d}x_1\mathrm{d}x_2 - \int_{x_1}\int_{x_2} f^\beta(x_1,x_2)\,\mathrm{d}x_1\mathrm{d}x_2\right] \\
&= \frac{1}{\beta-\alpha}\left[\left(1 - \int_{x_1}\int_{x_2} f^\beta(x_1,x_2)\,\mathrm{d}x_1\mathrm{d}x_2\right)\right. \\
&\qquad\qquad \left. - \left(1 - \int_{x_1}\int_{x_2} f^\alpha(x_1,x_2)\,\mathrm{d}x_1\mathrm{d}x_2\right)\right].
\end{aligned} \tag{2.11}$$

Equation (2.11) together with (2.10) reduces to

$$\begin{aligned}
H_{\alpha,\beta}(f) &= \frac{\beta-1}{\beta-\alpha}\left[H_\beta(f) + E_{f_1} f_1^{\beta-1} H_\beta(f_2|f_1)\right] \\
&\qquad - \frac{\beta-1}{\beta-\alpha}\left[H_\alpha(f) + E_{f_1} f_1^{\alpha-1} H_\alpha(f_2|f_1)\right], \\
\alpha \neq \beta, &\qquad 0<\alpha\le 1,\ \beta>1 \quad \text{or} \quad 0<\beta\le 1,\ \alpha>1.
\end{aligned} \tag{2.12}$$

3. Evaluation of β-class and (α, β)-class entropies

In this section, we consider $f(x)$ as multivariate normal density function with μ and Σ as mean vector and covariance matrix respectively. The choice of $f(x)$ in this form brings out the importance of β-class and (α, β)-class entropies which is established in the following two theorems:

Theorem 3.1. *The β-class entropy measure defined by*

$$H_\beta(f) = \frac{1}{\beta - 1}\left[1 - \int_x f^\beta(x)\,\mathrm{d}x\right], \qquad \beta \neq 1,\ \ \beta > 0, \tag{3.1}$$

contains more information than that of Shannon's entropy in case the density function $f(x)$ is given and $\beta < 1$.

P r o o f. Let us consider $f(x)$ to be a multivariate normal density function with mean vector μ and covariance matrix Σ. Then its pdf can be written as

$$f(x) = \frac{1}{(2\pi)^{\frac{n}{2}}|\Sigma|^{\frac{1}{2}}}\mathrm{e}^{-\frac{1}{2}(x-\mu)'\Sigma^{-1}(x-\mu)}, \qquad -\infty \le x \le \infty. \tag{3.2}$$

On putting $f(x)$ in (3.1), we get

$$\begin{aligned} H_\beta(f) &= \frac{1}{\beta - 1}\left[1 - \int_x \frac{1}{(2\pi)^{\frac{n\beta}{2}}|\Sigma|^{\frac{\beta}{2}}}\mathrm{e}^{-\frac{\beta}{2}(x-\mu)'\Sigma^{-1}(x-\mu)}\,\mathrm{d}x\right] \\ &= \frac{1}{\beta - 1}\left[1 - \frac{1}{(2\pi)^{\frac{n\beta}{2}}|\Sigma|^{\frac{\beta}{2}}}\int_x \mathrm{e}^{-\frac{\beta}{2}(x-\mu)'\Sigma^{-1}(x-\mu)}\,\mathrm{d}x\right]. \end{aligned} \tag{3.3}$$

Since $\frac{1}{(2\pi)^{\frac{n}{2}}|\Sigma|^{\frac{1}{2}}}\int_x \mathrm{e}^{-\frac{1}{2}(x-\mu)'\Sigma^{-1}(x-\mu)}\,\mathrm{d}x = 1$, (3.3) reduces to

$$\begin{aligned} H_\beta(f) &= \frac{1}{\beta - 1}\left[1 - (2\pi)^{\frac{-n\beta}{2}}|\Sigma|^{\frac{-\beta}{2}}(2\pi)^{\frac{n}{2}}|\Sigma|^{\frac{1}{2}}\beta^{\frac{-n}{2}}\right] \\ &= \frac{1}{\beta - 1}\left[1 - (2\pi)^{\frac{-n(\beta-1)}{2}}|\Sigma|^{\frac{-(\beta-1)}{2}}(\beta)^{\frac{-n}{2}}\right] \end{aligned} \tag{3.4}$$

which is the average information contained in the normal random vector X.

In case $f(x)$ is univariate exponential distribution with pdf

$$f(x) = \frac{1}{\theta}\mathrm{e}^{-\frac{x}{\theta}}, \qquad x > 0,\ \ \theta > 0, \tag{3.5}$$

then the average information in the exponential variable is

$$\begin{aligned} H_\beta(f) &= \frac{1}{\beta - 1}\left[1 - \int_x \frac{1}{\theta^\beta}\mathrm{e}^{-\frac{\beta x}{\theta}}\,\mathrm{d}x\right] \\ &= \frac{1}{\beta - 1}\left[1 - \frac{\theta^{1-\beta}}{\beta}\right]. \end{aligned} \tag{3.6}$$

It is worth mentioning that $H_\beta(f)$ changes with the choice of the entropy measure. $\square$

Next, we obtain the analogous results for pdf $f(x)$ by using Shannon's entropy as given below:

$$H(f) = -\int_x f(x)\log f(x)\,\mathrm{d}x. \tag{3.7}$$

In case the given pdf is univariate normal, then (3.7) gives

$$\begin{aligned} H(f) &= -\int_x \frac{1}{\sqrt{2\pi\sigma^2}}\mathrm{e}^{-\frac{1}{2}\left(\frac{x-\mu}{\sigma}\right)^2}\log\frac{1}{\sqrt{2\pi\sigma^2}}\mathrm{e}^{-\frac{1}{2}\left(\frac{x-\mu}{\sigma}\right)^2}\,\mathrm{d}x \\ &= -\int_x \left[-\frac{1}{2}\log 2\pi - \log\sigma - \frac{1}{2}\Big(\frac{x-\mu}{\sigma}\Big)^2\right]\frac{1}{\sqrt{2\pi\sigma^2}}\mathrm{e}^{-\frac{1}{2}\left(\frac{x-\mu}{\sigma}\right)^2}\,\mathrm{d}x \\ &= \frac{1}{2}\log 2\pi + \log\sigma + \frac{1}{2}. \end{aligned} \tag{3.8}$$

When $f(x)$ is exponentially distributed with parameter θ, we have

$$\begin{aligned} H(f) &= -\int_x \frac{1}{\theta}\mathrm{e}^{-\frac{x}{\theta}}\log\frac{1}{\theta}\mathrm{e}^{-\frac{x}{\theta}}\,\mathrm{d}x \\ &= -\int_x \frac{1}{\theta}\mathrm{e}^{-\frac{x}{\theta}}\left[-\log\theta - \frac{x}{\theta}\right]\mathrm{d}x \\ &= \log\theta + 1. \end{aligned} \tag{3.9}$$

$H_\beta(f)$ obtained in (3.6) for exponential distribution is function of θ and β. In case $\theta = 1$, then $H_\beta(f) = \frac{1}{\beta}$. Thus, when $\beta \geq 1$, Shannon's entropy gives maximum information and for $0 < \beta < 1$, β-class entropy contains more information about $f(x)$ than Shannon's entropy. Hence Theorem 3.1 is proved.

Theorem 3.2. *For a given density function $f(x)$, (α,β)-class entropy defined as*

$$H_{\alpha,\beta}(f) = \frac{1}{\beta-\alpha}\left[\int_x f^\alpha(x)\,\mathrm{d}x - \int_x f^\beta(x)\,\mathrm{d}x\right], \qquad \alpha \neq \beta;\ \ \alpha,\beta > 0 \tag{3.10}$$

contains more information than the information given by Shannon's entropy and β-class entropy.

Proof. Let $f(x)$ be the multivariate normal density function with mean vector μ and covariance matrix Σ. Then pdf $f(x)$ is given by

$$f(x) = \frac{1}{(2\pi)^{\frac{n}{2}}|\Sigma|^{\frac{1}{2}}}\mathrm{e}^{-\frac{1}{2}(x-\mu)'\Sigma^{-1}(x-\mu)}, \qquad -\infty \leq x \leq \infty. \tag{3.11}$$

On substituting (3.11) in (3.10) and simplification, we have

$$\begin{aligned} H_{\alpha,\beta}(f) = \frac{1}{\beta-\alpha}(2\pi)^{\frac{n}{2}}|\Sigma|^{\frac{1}{2}}\big[(2\pi)^{-\frac{n\alpha}{2}}|\Sigma|^{-\frac{\alpha}{2}}\alpha^{-\frac{n}{2}} \\ - (2\pi)^{-\frac{n\beta}{2}}|\Sigma|^{-\frac{\beta}{2}}\beta^{-\frac{n}{2}}\big], \end{aligned} \tag{3.12}$$

which is the average information contained in the normal random vector X. □

Let us consider the particular case of $f(x)$ when it is univariate i.e. $n = 1$ and $\Sigma = \sigma^2 I$. In this case (3.12) reduces to

$$\begin{aligned} H_{\alpha,\beta}(f) &= \frac{1}{\beta-\alpha}(2\pi)^{\frac{1}{2}}|\sigma^2 I|^{\frac{1}{2}}\Big[(2\pi)^{-\frac{\alpha}{2}}|\sigma^2 I|^{-\frac{\alpha}{2}}\alpha^{-\frac{1}{2}} - (2\pi)^{-\frac{\beta}{2}}|\sigma^2 I|^{-\frac{\beta}{2}}\beta^{-\frac{1}{2}}\Big] \\ &= \frac{1}{\beta-\alpha}\sigma\sqrt{2\pi}\Big[(2\pi)^{-\frac{\alpha}{2}}\sigma^{-\alpha}\alpha^{-\frac{1}{2}} - (2\pi)^{-\frac{\beta}{2}}\sigma^{-\beta}\beta^{-\frac{1}{2}}\Big]. \end{aligned} \tag{3.13}$$

When $f(x)$ is univariate exponential distribution with pdf

$$f(x) = \frac{1}{\theta}\mathrm{e}^{-\frac{x}{\theta}}, \qquad x > 0, \ \ \theta > 0,$$

then (3.10) becomes

$$\begin{aligned} H_{\alpha,\beta}(f) &= \frac{1}{\beta-\alpha}\bigg[\int_x \Big(\frac{1}{\theta}\mathrm{e}^{-\frac{x}{\theta}}\Big)^{\alpha}\mathrm{d}x - \int_x \Big(\frac{1}{\theta}\mathrm{e}^{-\frac{x}{\theta}}\Big)^{\beta}\mathrm{d}x\bigg] \\ &= \frac{1}{\beta-\alpha}\bigg[\frac{\theta^{-\alpha+1}}{\alpha} - \frac{\theta^{-\beta+1}}{\beta}\bigg], \end{aligned} \tag{3.14}$$

which is the average information contained in the exponentially distributed variable X.

Particular Cases:

Case 1. When $\alpha \to 1$, equation (3.13) reduces to

$$H_{1,\beta}(f) = \frac{1}{\beta-1}\Big[1 - (2\pi)^{-\frac{(\beta-1)}{2}}\sigma^{-(\beta-1)}\beta^{-\frac{1}{2}}\Big], \tag{3.15}$$

which is the same given by β-class entropies when $f(x)$ is univariate normal density function. Similarly, in case $\alpha \to 1$, (3.14) becomes

$$H_{1,\beta}(f) = \frac{1}{\beta-1}\bigg[1 - \frac{\theta^{-(\beta-1)}}{\beta}\bigg], \tag{3.16}$$

which is (3.6).

Case 2. In case $\beta \to 1$, (3.15) reduces to

$$H_{1,1}(f) = \frac{1}{2} + \log\sqrt{2\pi} + \log\sigma, \tag{3.17}$$

which is (3.8). Similarly, when $\beta \to 1$, equation (3.16) reduces to (3.9).

If we take $\theta = 1$ in (3.14), then it becomes

$$H_{\alpha,\beta}(f) = \frac{1}{\beta-\alpha}\left[\frac{1}{\alpha} - \frac{1}{\beta}\right] = \frac{1}{\alpha\beta}. \tag{3.18}$$

It can be inferred from (3.17) that when $0 < \alpha < 1$ and $1 < \beta < \frac{1}{\alpha}$, then $H_{\alpha,\beta}(f)$ contains more information than Shannon's entropy and α-class entropies and in case $\alpha = 1$, it contains the same amount of information as β-class contains. Similarly, when $0 < \beta < 1$ and $1 < \alpha < \frac{1}{\beta}$, (3.17) contains more information than Shannon's entropy and β-class entropy and in case $\beta = 1$, then it contains the same amount of information as α-class entropy contains. Thus, we can conclude that the parametric generalized measures of entropy for a given normal distribution are more advantageous as these give more information about the density function than Shannon's entropy with appropriate choice of parameter.

4. Generalized measures of dependence

Let $f = f(x) = f[x_1, x_2]$ be the joint density function $f(x_1) = f_1$ and $f(x_2) = f_2$ be the marginal density functions of x_1 and x_2 respectively. In [7], the following measure of dependence between x_1 and x_2 has been defined:

$$\begin{aligned} D(f : f_1 f_2) &= H(f) - H(f_1 f_2) \\ &= H(f) - \big[H(f_1) + H(f_2)\big], \end{aligned} \tag{4.1}$$

where $H(\cdot)$ is Shannon's entropy.

(4.1) can be generalized in many ways, but here we shall consider the following three forms only:

(a) Let

$$D_\alpha(f : f_1 f_2) = H_\alpha(f) - H_\alpha(f_1 f_2), \tag{4.2}$$

where

$$H_\alpha(f) = \frac{1}{1-\alpha} \log \int_x f^\alpha(x)\,\mathrm{d}x$$

and

$$H_\alpha(f_1 f_2) = \frac{1}{1-\alpha} \log \int_{x_1}\int_{x_2} f^\alpha(x_1) f^\alpha(x_2)\,\mathrm{d}x_1 \mathrm{d}x_2 = H_\alpha(f_1) + H_\alpha(f_2).$$

Then (4.2) reduces to

$$D_\alpha(f : f_1 f_2) = H_\alpha(f) - \big[H_\alpha(f_1) + H_\alpha(f_2)\big]$$

(b) Let

$$D_\beta(f : f_1 f_2) = H_\beta(f) - H_\beta(f_1 f_2), \tag{4.3}$$

where

$$H_\beta(f) = \frac{1}{1-\beta}\Big[1 - \int_x f^\beta(x)\,\mathrm{d}x\Big]$$

and

$$H_\beta(f_1 f_2) = \frac{1}{1-\beta}\Big[1 - \int_{x_1}\int_{x_2} f^\beta(x_1) f^\beta(x_2)\,\mathrm{d}x_1 \mathrm{d}x_2\Big]$$
$$= H_\beta(f_1) + \int_{x_1} f^\beta(x_1) H_\beta(f_2)\,\mathrm{d}x_1\Big]$$

Then

$$D_\beta(f : f_1 f_2) = H_\beta(f) - \Big[H_\beta(f_1) + \int_{x_1} f_1^\beta H_\beta(f_2)\,\mathrm{d}x_1\Big].$$

(c) Let

$$D_{\alpha,\beta}(f : f_1 f_2) = H_{\alpha,\beta}(f) - H_{\alpha,\beta}(f_1 f_2), \tag{4.4}$$

where

$$H_{\alpha,\beta}(f) = \frac{1}{\beta-\alpha}\Big[\int_x f^\alpha(x)\,\mathrm{d}x - \int_x f^\beta(x)\,\mathrm{d}x\Big],$$

$$H_{\alpha,\beta}(f_1 f_2) =$$
$$= \frac{1}{\beta-\alpha}\Big[\int_{x_1}\int_{x_2} (f_1 f_2)^\alpha\,\mathrm{d}x_1\mathrm{d}x_2 - \int_{x_1}\int_{x_2} (f_1 f_2)^\beta\,\mathrm{d}x_1\mathrm{d}x_2\Big]$$
$$= \frac{1-\beta}{\beta-\alpha}\Big[H_\beta(f_1) + \int_{x_1} f_1^\beta H_\beta(f_2)\,\mathrm{d}x_1\Big] - \frac{1-\alpha}{\beta-\alpha}\Big[H_\alpha(f_1) + \int_{x_1} f_1^\alpha H_\alpha(f_2)\,\mathrm{d}x_1\Big].$$

Then

$$D_{\alpha,\beta}(f : f_1 f_2) = H_{\alpha,\beta}(f) - \frac{1-\beta}{\beta-\alpha}\Big[H_\beta(f_1) + \int_{x_1} f_1^\beta H_\beta(f_2)\,\mathrm{d}x_1\Big]$$
$$+ \frac{1-\alpha}{\beta-\alpha}\Big[H_\alpha(f_1) + \int_{x_1} f_1^\alpha H_\alpha(f_2)\,\mathrm{d}x_1\Big].$$

All the measures given by (4.2), (4.3) and (4.4) can be described as the measures of mutual information in the sense that they measure the amount of information x_2 contains about x_1.

In case $\alpha = 1$ in (4.2), $\beta = 1$ in (4.3) and $\alpha = \beta = 1$ in (4.4), these reduce to

$$D_1(f : f_1 f_2) = H(f) - \big[H(f_1) + H(f_2)\big],$$

which is (4.1). x_1 and x_2 are independent iff $D_{\alpha,\beta} = D_\alpha = D_\beta = D = 0$.

REFERENCES

[1] GOKHALE, D. V.: *On joint conditional entropies*, Entropy **1** (1999), 21–24.
[2] HARVDA, J.—CHARVÁT, F.: *Quantification method of classification processes-concept of structural β-entropy*, Kybernetika (Prague) **3** (1967), 30–35.
[3] KAPUR, J. N.—KESAVAN, H. K.: *Entropy Optimization Principles with Applications*, Academic Press, San Diego, 1992.
[4] RÉNYI A.: *On measures of entropy and information*, Proc. 4th Berkeley Symp. Math. Stat. Probab. **1** (1961), 546–561.
[5] SHANNON, C. E.: *A mathematical theory of communication*, Bell System Tech. J. **27** (1948), 379–423.
[6] SHARMA, B. D.—TANEJA, I. J.: *Three generalized additive measures of entropy*, Elektronische Information Verarbeitung und Kybernetik **13** (1977), 419–433.
[7] ULLAH, A.: *Entropy, divergence and distance measures with econometric applications*, J. Statist. Plann. Inference **49** (1996), 137–162.

Received 11.8.2005
Revised 14.3.2007

* *Department of Statistics*
University of Jammu
Jammu-180 004
INDIA

** *Jaypee Institute of Engineering and Technology*
A.B.Road, Raghogarh
Distt. Guna-473226
INDIA