

DIRECTED DIVERGENCE AS A MEASURE OF SIMILARITY

Om Parkash¹, Mukesh² §

^{1,2}Department of Mathematics

Guru Nanak Dev University

Amritsar, 143005, INDIA

Abstract: The development of interrelationships between divergence measures and the known statistical constants provide the applications of information theory to the field of statistics. In the literature of information measures, there exist many divergence measures for discrete probability distributions whereas we need such divergence measures for continuous distributions to extend the scope of their applications. In the present communication, we have introduced divergence measures for continuous variate distributions and then proved that the divergence between the joint distribution density and the product of the marginal distribution densities is a function of the correlation coefficient which obviously implies that the divergence is also a measure of the similarity or of the dissimilarity.

AMS Subject Classification: 94A15, 94A17

Key Words: Gaussian distribution, divergence, Chi-square statistic, correlation coefficient

1. Introduction

It has been observed that the two fundamental concepts, namely, entropy and relative entropy or divergence measure play a significant role in the field of information theory because of their wide applicability towards engineering and other applied sciences. In particular, the measure of relative entropy may be more fundamental and appropriate while dealing with the problems concerning statistical analysis. Joe [4] discussed and investigated the role of measures

Received: October 24, 2013

© 2014 Academic Publications, Ltd.
url: www.acadpubl.eu

§Correspondence author

of multivariate dependence and measures of conditional dependence based on relative entropies and emphasized that for continuous variables, a certain transformation of relative entropy to the interval $[0, 1]$ leads to generalizations of the correlation, multiple-correlation and partial-correlation coefficients. In the literature of statistics, there are some statistics developed to measure the degree of association between qualitative variables. These statistics are mainly derived from the chi-square value calculated for a contingency table. Besides, some statistics based on entropy measures are widely used to measure qualitative association. Evren and Tuna [2] pointed out that the statistics derived from the concept of entropy like mutual information and Kullback-Leibler's [6] divergence measure can also be used in measuring the degree of association between qualitative variables.

Hafidia and Mkhadria [3] investigated and evaluated an extension of the Akaike [1] information criterion, which is an approximately unbiased estimator for a risk function based on the Kullback's [5] symmetric divergence. Toussaint [14] introduced a divergence measure between populations as a generalization of the J-divergence. Menéndez et al. [7] obtained the asymptotic distribution of Toussaint's [14] measure when the parameters are substituted by their maximum likelihood estimators and the probability density functions belong to the exponential family. Parkash and Thukral [13] emphasized that the two fields of research, namely, statistics and information theory have found tremendous applications in the analysis of biological data. Statistics is extensively used for the measurement of statistical constants whereas measures of information are used to study diversity and equitability. Keeping in view the idea, the authors developed the link between the two and proved that statistical constants can be used as information measures.

In statistics, the entropy corresponds to the maximum likelihood method, in which Kullback-Leibler [6] divergence measure, given by

$$D_{KL}(P; Q) = \sum_{i=1}^n p_i \log \frac{p_i}{q_i}, \quad (1)$$

connects Boltzmann-Shannon entropy and the expected log-likelihood function.

Keeping in view the applicability of directed divergence to the various fields of science and technology, Parkash and Mukesh [12] have developed a new symmetric directed divergence for the discrete probability distributions P and Q given by

$$D(P; Q) = \sum_{i=1}^n \left(\frac{p_i^2}{q_i} + \frac{q_i^2}{p_i} - 2p_i \right). \quad (2)$$

Further, Parkash and Mukesh [8] developed a new generalized parametric measure of divergence for the probability distributions P and Q which is given by the following mathematical expression

$$D_{\alpha,\beta}(P; Q) = \frac{\sum_{i=1}^n p_i^{\alpha-\beta+1} q_i^{\beta-\alpha} - 1}{\alpha - \beta}, \alpha \neq \beta, \beta < \alpha + 1, -\infty < \alpha < \infty. \quad (3)$$

Parkash and Mukesh [11] investigated certain relations between parametric information measures and chi-square statistic by using the optimization principles. These relations were developed using parametric measures of information introduced recently by Parkash and Mukesh [10]. Further, Parkash and Mukesh [9] provided the applications of divergence measure by developing an optimizational principle for minimizing risk in portfolio analysis.

In the present communication, we have extended our results to continuous variate distributions to prove that the divergence between the joint distribution density and the product of the marginal distribution densities is a function of the correlation coefficient.

2. New Measure of Divergence Between the Joint Distribution Density and the Product of the Marginal Distribution Densities

In this section, we make the use of measure of directed divergence (1.2) to find the distance between the joint distribution density and the product of the marginal distribution densities for Gaussian distributed random variables.

We consider the continuous version of the divergence (1.2) between the probability distributions with density functions $f_{\underline{x}}(\underline{\xi})$ and $f_{\underline{x}_s}(\underline{\xi})$, which is given by

$$D(f_{\underline{x}}(\underline{\xi}), f_{\underline{x}_s}(\underline{\xi})) = \int_{\underline{\Xi}} \left[\frac{f_{\underline{x}}^2(\underline{\xi})}{f_{\underline{x}_s}(\underline{\xi})} + \frac{f_{\underline{x}_s}^2(\underline{\xi})}{f_{\underline{x}}(\underline{\xi})} - 2f_{\underline{x}}(\underline{\xi}) \right] d\underline{\xi}. \quad (4)$$

We start with the examination of the special case of two Gaussian variables x_1 and x_2 , which is summarized in one vector \underline{x} , given by

$$f_{\underline{x}}(\underline{\xi}) = \frac{1}{(2\pi)^{|\Sigma_{xx}|^{\frac{1}{2}}}} \exp\left[-\frac{1}{2}(\underline{\xi} - \underline{\mu}_{\underline{x}})^T \cdot \Sigma_{xx}^{-1} \cdot (\underline{\xi} - \underline{\mu}_{\underline{x}})\right], \quad (5)$$

where $\underline{x}^T = [x_1, x_2]$, $\underline{\xi}^T = [\xi_1, \xi_2]$, $\underline{\mu}_{\underline{x}}^T = [\mu_1^T, \mu_2^T]$ and $\Sigma_{xx} = \begin{bmatrix} \sigma_1^2 & \rho\sigma_1\sigma_2 \\ \rho\sigma_1\sigma_2 & \sigma_2^2 \end{bmatrix}$.

We obtain the determinant of the covariance matrix

$$|\Sigma_{xx}| = \begin{vmatrix} \sigma_1^2 & \rho\sigma_1\sigma_2 \\ \rho\sigma_1\sigma_2 & \sigma_2^2 \end{vmatrix} = \sigma_1^2\sigma_2^2(1 - \rho^2)$$

and the inverse matrix

$$\Sigma_{xx}^{-1} = \frac{1}{\sigma_1^2 \sigma_2^2 (1 - \rho^2)} \begin{bmatrix} \sigma_2^2 & -\rho \sigma_1 \sigma_2 \\ -\rho \sigma_1 \sigma_2 & \sigma_1^2 \end{bmatrix} \tag{6}$$

Using equation (2.3), the quadratic form in the exponential function of equation (2.2) becomes

$$(\underline{\xi} - \underline{\mu}_{\underline{x}})^T \cdot \Sigma_{xx}^{-1} \cdot (\underline{\xi} - \underline{\mu}_{\underline{x}}) = \frac{(\xi_1 - \mu_1)^2}{\sigma_1^2 (1 - \rho^2)} - \frac{2\rho(\xi_1 - \mu_1)(\xi_2 - \mu_2)}{\sigma_1 \sigma_2 (1 - \rho^2)} + \frac{(\xi_2 - \mu_2)^2}{\sigma_2^2 (1 - \rho^2)}. \tag{7}$$

Now the divergence between the joint distribution density $f_{\underline{x}}(\underline{\xi})$ and the product of the marginal distribution densities $f_{x_1}(\xi_1) \cdot f_{x_2}(\xi_2)$ is given by

$$D(f_{\underline{x}}(\underline{\xi}), f_{x_1}(\xi_1) \cdot f_{x_2}(\xi_2)) = \int_{\Xi} \left[\frac{f_{\underline{x}}^2(\underline{\xi})}{f_{x_1}(\xi_1) \cdot f_{x_2}(\xi_2)} + \frac{f_{x_1}^2(\xi_1) \cdot f_{x_2}^2(\xi_2)}{f_{\underline{x}}(\underline{\xi})} - 2f_{\underline{x}}(\underline{\xi}) \right] d\underline{\xi}. \tag{8}$$

If the random variables are independent, the divergence vanishes and we get $D = 0$, because in this case the joint distribution density function is equal to the product of the marginal distribution functions.

Thus

$$D = \int_{\Xi} \frac{f_{\underline{x}}^2(\underline{\xi})}{f_{x_1}(\xi_1) \cdot f_{x_2}(\xi_2)} d\underline{\xi} + \int_{\Xi} \frac{f_{x_1}^2(\xi_1) \cdot f_{x_2}^2(\xi_2)}{f_{\underline{x}}(\underline{\xi})} d\underline{\xi} - 2. \tag{9}$$

Now

$$\begin{aligned} \frac{f_{\underline{x}}^2(\underline{\xi})}{f_{x_1}(\xi_1) \cdot f_{x_2}(\xi_2)} &= \frac{\frac{1}{4\pi^2 \sigma_1^2 \sigma_2^2 (1 - \rho^2)} \exp\left[-\frac{(\xi_1 - \mu_1)^2}{\sigma_1^2 (1 - \rho^2)} + \frac{2\rho(\xi_1 - \mu_1)(\xi_2 - \mu_2)}{\sigma_1 \sigma_2 (1 - \rho^2)} - \frac{(\xi_2 - \mu_2)^2}{\sigma_2^2 (1 - \rho^2)}\right]}{\frac{1}{\sqrt{2\pi}\sigma_1} \exp\left[-\frac{(\xi_1 - \mu_1)^2}{2\sigma_1^2}\right] \cdot \frac{1}{\sqrt{2\pi}\sigma_2} \exp\left[-\frac{(\xi_2 - \mu_2)^2}{2\sigma_2^2}\right]} \\ &= \frac{\exp\left[-\frac{(1 + \rho^2)(\xi_1 - \mu_1)^2}{2\sigma_1^2 (1 - \rho^2)} + \frac{2\rho(\xi_1 - \mu_1)(\xi_2 - \mu_2)}{\sigma_1 \sigma_2 (1 - \rho^2)} - \frac{(1 + \rho^2)(\xi_2 - \mu_2)^2}{2\sigma_2^2 (1 - \rho^2)}\right]}{(2\pi)\sigma_1 \sigma_2 (1 - \rho^2)}. \end{aligned}$$

Therefore

$$\int_{\Xi} \frac{f_{\underline{x}}^2(\underline{\xi})}{f_{x_1}(\xi_1) \cdot f_{x_2}(\xi_2)} d\underline{\xi} = \frac{I}{(2\pi)\sigma_1 \sigma_2 (1 - \rho^2)} \tag{10}$$

where

$$\begin{aligned} I &= \int_{\Xi} \exp\left[-\frac{(1 + \rho^2)(\xi_1 - \mu_1)^2}{2\sigma_1^2 (1 - \rho^2)} + \frac{2\rho(\xi_1 - \mu_1)(\xi_2 - \mu_2)}{\sigma_1 \sigma_2 (1 - \rho^2)} - \frac{(1 + \rho^2)(\xi_2 - \mu_2)^2}{2\sigma_2^2 (1 - \rho^2)}\right] d\underline{\xi} \\ &= \int_{\Xi_2} \exp\left[-\frac{(1 + \rho^2)(\xi_2 - \mu_2)^2}{2\sigma_2^2 (1 - \rho^2)}\right] \left\{ \int_{\Xi_1} \exp\left[-\frac{(1 + \rho^2)(\xi_1 - \mu_1)^2}{2\sigma_1^2 (1 - \rho^2)} + \frac{2\rho(\xi_1 - \mu_1)(\xi_2 - \mu_2)}{\sigma_1 \sigma_2 (1 - \rho^2)}\right] d\xi_1 \right\} d\xi_2 \\ &= \int_{\Xi_2} \exp\left[-\frac{(1 + \rho^2)(\xi_2 - \mu_2)^2}{2\sigma_2^2 (1 - \rho^2)}\right] \left\{ \sqrt{\frac{2\pi\sigma_1^2 (1 - \rho^2)}{(1 + \rho^2)}} \exp\left[\frac{2\rho^2(\xi_2 - \mu_2)^2}{\sigma_1^2 (1 - \rho^2)(1 + \rho^2)}\right] \right\} d\xi_2 \end{aligned}$$

$$\begin{aligned}
 &= \sqrt{\frac{2\pi\sigma_1^2(1-\rho^2)}{(1+\rho^2)}} \int_{\Xi_2} \exp\left[-\frac{(1-\rho^2)(\xi_2-\mu_2)^2}{2\sigma_2^2(1+\rho^2)}\right] d\xi_2 \\
 &= \sqrt{\frac{2\pi\sigma_1^2(1-\rho^2)}{(1+\rho^2)}} \left[\sqrt{\frac{2\pi\sigma_2^2(1+\rho^2)}{(1-\rho^2)}} \right] \\
 &= 2\pi\sigma_1\sigma_2,
 \end{aligned}$$

which is used in equation (2.7), to give

$$\int_{\Xi} \frac{f_{\underline{x}}^2(\underline{\xi})}{f_{x_1}(\xi_1) \cdot f_{x_2}(\xi_2)} d\underline{\xi} = \frac{1}{(1-\rho^2)}. \tag{11}$$

Again

$$\begin{aligned}
 \frac{f_{x_1}^2(\xi_1) \cdot f_{x_2}^2(\xi_2)}{f_{\underline{x}}(\underline{\xi})} &= \frac{\sqrt{1-\rho^2}}{(2\pi)\sigma_1\sigma_2} \exp\left[-\frac{(1-2\rho^2)(\xi_1-\mu_1)^2}{2\sigma_1^2(1-\rho^2)} - \frac{\rho(\xi_1-\mu_1)(\xi_2-\mu_2)}{\sigma_1\sigma_2(1-\rho^2)} \right. \\
 &\quad \left. - \frac{(1-2\rho^2)(\xi_2-\mu_2)^2}{2\sigma_2^2(1-\rho^2)}\right].
 \end{aligned}$$

Therefore

$$\int_{\Xi} \frac{f_{x_1}^2(\xi_1) \cdot f_{x_2}^2(\xi_2)}{f_{\underline{x}}(\underline{\xi})} d\underline{\xi} = \frac{\sqrt{1-\rho^2}}{(2\pi)\sigma_1\sigma_2} (I') \tag{12}$$

where

$$\begin{aligned}
 I' &= \int_{\Xi} \exp\left[-\frac{(1-2\rho^2)(\xi_1-\mu_1)^2}{2\sigma_1^2(1-\rho^2)} - \frac{\rho(\xi_1-\mu_1)(\xi_2-\mu_2)}{\sigma_1\sigma_2(1-\rho^2)} - \frac{(1-2\rho^2)(\xi_2-\mu_2)^2}{2\sigma_2^2(1-\rho^2)}\right] d\underline{\xi} \\
 &= \int_{\Xi_2} \exp\left[-\frac{(1-2\rho^2)(\xi_2-\mu_2)^2}{2\sigma_2^2(1-\rho^2)}\right] \left\{ \int_{\Xi_1} \exp\left[-\frac{(1-2\rho^2)(\xi_1-\mu_1)^2}{2\sigma_1^2(1-\rho^2)} \right. \right. \\
 &\quad \left. \left. - \frac{\rho(\xi_1-\mu_1)(\xi_2-\mu_2)}{\sigma_1\sigma_2(1-\rho^2)}\right] d\xi_1 \right\} d\xi_2 \\
 &= \int_{\Xi_2} \exp\left[-\frac{(1-2\rho^2)(\xi_2-\mu_2)^2}{2\sigma_2^2(1-\rho^2)}\right] \left\{ \sqrt{\frac{2\pi\sigma_1^2(1-\rho^2)}{1-2\rho^2}} \exp\left[\frac{\rho^2(\xi_2-\mu_2)^2}{2\sigma_2^2(1-\rho^2)(1-2\rho^2)}\right] \right\} d\xi_2
 \end{aligned}$$

provided $-\frac{1}{\sqrt{2}} < \rho < \frac{1}{\sqrt{2}}$. Thus equation (2.9) becomes

$$\begin{aligned}
 \int_{\Xi} \frac{f_{x_1}^2(\xi_1) \cdot f_{x_2}^2(\xi_2)}{f_{\underline{x}}(\underline{\xi})} d\underline{\xi} &= \frac{(1-\rho^2)}{\sqrt{2\pi}\sigma_2\sqrt{1-2\rho^2}} \int_{\Xi_2} \exp\left[-\frac{(1+4\rho^4-5\rho^2)(\xi_2-\mu_2)^2}{2\sigma_2^2(1-\rho^2)(1-2\rho^2)}\right] d\xi_2 \\
 &= \frac{(1-\rho^2)}{\sqrt{2\pi}\sigma_2\sqrt{1-2\rho^2}} \left[\sqrt{\frac{2\pi\sigma_2^2(1-2\rho^2)}{(1-4\rho^2)}} \right]
 \end{aligned}$$

provided $-\frac{1}{2} < \rho < \frac{1}{2}$. Therefore

$$\int_{\Xi} \frac{f_{x_1}^2(\xi_1) \cdot f_{x_2}^2(\xi_2)}{f_{\underline{x}}(\underline{\xi})} d\underline{\xi} = \frac{(1-\rho^2)}{\sqrt{1-4\rho^2}}. \tag{13}$$

Using equations (2.8) and (2.10) in equation (2.6), we get

$$D = \frac{1}{(1 - \rho^2)} + \frac{(1 - \rho^2)}{\sqrt{1 - 4\rho^2}} - 2, \quad (14)$$

provided $-\frac{1}{2} < \rho < \frac{1}{2}$.

Thus, the divergence (2.5) between the joint distribution density and the product of the marginal distribution densities is a function of the correlation coefficient. So the divergence is also a measure of the similarity or of the dissimilarity. The existing diversity of the two distribution densities examined arises because on the one hand we generate the joint distribution as a product of the marginal distribution densities of independent random variables and on the other hand we have the joint distribution density of correlated random variables. The difference between these distribution densities obviously lies in the correlation, which occurs as parameter in the result of our computation.

Further, at the extremes we have following possibilities:

When $\rho = 0$, the two random variables are not correlated and thus the divergence vanishes, because in this case the joint distribution density is equal to the product of the marginal distribution densities.

When $\rho = 1$, the divergence achieves an infinitely large value, because both random variables are identical. Thus there actually is no joint distribution density anymore, because one of the two random variables is generated by a deterministic mapping from the other.

3. New Generalized Measure of Divergence between The Joint Distribution Density and the Product of The Marginal Distribution Densities

In this section, we have extended the results of the non-parametric measure discussed in Section 2 to parametric one. To find the distance between the joint distribution density and the product of the marginal distribution densities for Gaussian distributed random variables, we have restricted the parametric range of the generalized measure of directed divergence (1.3).

We consider the continuous version of the generalized measure of divergence (1.3) with the restricted range of parameters. This divergence for the continuous-variate probability distributions with new range of parameters for the density functions $f_{\underline{x}}(\underline{\xi})$ and $f_{\underline{x}_s}(\underline{\xi})$ is given by

$$D_{\alpha, \beta}(f_{\underline{x}}(\underline{\xi}), f_{\underline{x}_s}(\underline{\xi})) = \frac{\int_{\underline{\Xi}} f_{\underline{x}}^{\alpha - \beta + 1}(\underline{\xi}) f_{\underline{x}_s}^{\beta - \alpha}(\underline{\xi}) d\underline{\xi} - 1}{\alpha - \beta}, \alpha \neq \beta, \beta < \alpha, -\infty < \alpha < \infty. \quad (15)$$

The divergence between the joint distribution density $f_{\underline{x}}(\underline{\xi})$ and the product of the marginal distribution densities $f_{x_1}(\xi_1) \cdot f_{x_2}(\xi_2)$ of two Gaussian variables x_1 and x_2 , summarized in one vector \underline{x} , is given by

$$D_{\alpha,\beta}(f_{\underline{x}}(\underline{\xi}), f_{x_1}(\xi_1) \cdot f_{x_2}(\xi_2)) = \frac{\int_{\Xi} f_{\underline{x}}^{\alpha-\beta+1}(\underline{\xi}) f_{x_1}^{\beta-\alpha}(\xi_1) \cdot f_{x_2}^{\beta-\alpha}(\xi_2) d\underline{\xi} - 1}{\alpha - \beta}. \quad (16)$$

Now the integrand

$$f_{\underline{x}}^{\alpha-\beta+1}(\underline{\xi}) f_{x_1}^{\beta-\alpha}(\xi_1) \cdot f_{x_2}^{\beta-\alpha}(\xi_2) = \frac{1}{2\pi\sigma_1\sigma_2(1-\rho^2)^{\frac{\alpha-\beta+1}{2}}} \exp\left\{-\frac{[1+(\alpha-\beta)\rho^2](\xi_1-\mu_1)^2}{2\sigma_1^2(1-\rho^2)} + \frac{(\alpha-\beta+1)\rho(\xi_1-\mu_1)(\xi_2-\mu_2)}{\sigma_1\sigma_2(1-\rho^2)} - \frac{[1+(\alpha-\beta)\rho^2](\xi_2-\mu_2)^2}{2\sigma_2^2(1-\rho^2)}\right\},$$

which is used in equation (3.2) to get

$$D_{\alpha,\beta}(f_{\underline{x}}(\underline{\xi}), f_{x_1}(\xi_1) \cdot f_{x_2}(\xi_2)) = \frac{J}{2\pi(\alpha-\beta)\sigma_1\sigma_2(1-\rho^2)^{\frac{\alpha-\beta+1}{2}}} - \frac{1}{(\alpha-\beta)}, \quad (17)$$

where

$$\begin{aligned} J &= \int_{\Xi} \exp\left\{-\frac{[1+(\alpha-\beta)\rho^2](\xi_1-\mu_1)^2}{2\sigma_1^2(1-\rho^2)} + \frac{(\alpha-\beta+1)\rho(\xi_1-\mu_1)(\xi_2-\mu_2)}{\sigma_1\sigma_2(1-\rho^2)} - \frac{[1+(\alpha-\beta)\rho^2](\xi_2-\mu_2)^2}{2\sigma_2^2(1-\rho^2)}\right\} d\underline{\xi} \\ &= \int_{\Xi_2} \exp\left\{-\frac{[1+(\alpha-\beta)\rho^2](\xi_2-\mu_2)^2}{2\sigma_2^2(1-\rho^2)}\right\} \left(\int_{\Xi_1} \exp\left\{-\frac{[1+(\alpha-\beta)\rho^2](\xi_1-\mu_1)^2}{2\sigma_1^2(1-\rho^2)} + \frac{(\alpha-\beta+1)\rho(\xi_1-\mu_1)(\xi_2-\mu_2)}{\sigma_1\sigma_2(1-\rho^2)}\right\} d\xi_1\right) d\xi_2 \\ &= \int_{\Xi_2} \exp\left\{-\frac{[1+(\alpha-\beta)\rho^2](\xi_2-\mu_2)^2}{2\sigma_2^2(1-\rho^2)}\right\} \left(\sqrt{\frac{2\pi\sigma_1^2(1-\rho^2)}{1+(\alpha-\beta)\rho^2}} \cdot \exp\left\{\frac{(\alpha-\beta+1)^2\rho^2(\xi_2-\mu_2)^2}{2\sigma_2^2(1-\rho^2)[1+(\alpha-\beta)\rho^2]}\right\}\right) d\xi_2 \\ &= \sqrt{\frac{2\pi\sigma_1^2(1-\rho^2)}{1+(\alpha-\beta)\rho^2}} \int_{\Xi_2} \exp\left\{-\frac{[1-(\alpha-\beta)^2\rho^2](\xi_2-\mu_2)^2}{2\sigma_2^2[1+(\alpha-\beta)\rho^2]}\right\} d\xi_2 \\ &= \sqrt{\frac{2\pi\sigma_1^2(1-\rho^2)}{1+(\alpha-\beta)\rho^2}} \sqrt{\frac{2\pi\sigma_2^2[1+(\alpha-\beta)\rho^2]}{1-(\alpha-\beta)^2\rho^2}} \\ &= \frac{2\pi\sigma_1\sigma_2\sqrt{1-\rho^2}}{\sqrt{1-(\alpha-\beta)^2\rho^2}} \end{aligned} \quad (18)$$

provided $\max\{-1, -\frac{1}{\alpha-\beta}\} < \rho < \min\{\frac{1}{\alpha-\beta}, 1\}$.

Using equation (3.4) in equation (3.3), we get

$$D_{\alpha,\beta} = \frac{1}{(\alpha - \beta)} \left[\frac{1}{(1 - \rho^2)^{\frac{\alpha - \beta}{2}} \sqrt{1 - (\alpha - \beta)^2 \rho^2}} - 1 \right], \quad (19)$$

with $\max\{-1, -\frac{1}{\alpha - \beta}\} < \rho < \min\{\frac{1}{\alpha - \beta}, 1\}$.

Thus, the divergence (3.2) between the joint distribution density and the product of the marginal distribution densities is a function of the correlation coefficient and thus it is a measure of the similarity or of the dissimilarity. Further, at the extremes we have following possibilities:

When $\rho = 0$, the two random variables are not correlated and thus the divergence vanishes.

When $\rho = 1$, the divergence achieves an infinitely large value, because both random variables are identical.

4. Concluding Remarks

We have introduced divergence measures for continuous variate distributions and then proved that this divergence is a function of the correlation coefficient which obviously implies that the divergence is also a measure of the similarity or of the dissimilarity and consequently, provides applications to the field of statistics. With similar arguments, one can provide the applications of divergence measures either by using other standards measures of divergence or by proving divergence to be a function of other statistical constants.

Acknowledgments

The authors are thankful to the University Grants Commission, New Delhi, for providing financial assistance for the preparation of the manuscript.

References

- [1] H. Akaike, Information theory and an extension of the maximum likelihood principle, In: B.N. Petrov and F. Csaki, Eds., Second International Symposium Information Theory, Akademia Kiado, Bubapest, (1973), 267-281.

- [2] A. Evren, E. Tuna, A comparison between entropy-based association measures and other qualitative association measures, *Selçuk Journal of Applied Mathematics*, **Special Issue** (2012), 3-17.
- [3] B. Hafidia, A. Mkhadria, An Akaike criterion based on Kullback symmetric divergence in the presence of incomplete-data, *Afrika Statistika*, **2** (2007), 1-21.
- [4] H. Joe, Relative entropy measures of multivariate dependence, *Journal of American Statistical Association*, **84** (1989), 157-164.
- [5] S. Kullback, *Information Theory and Statistics*, Wiley, New York, (1959).
- [6] S. Kullback, R.A. Leibler, On information and sufficiency, *Annals of Mathematical Statistics*, **22** (1951), 79-86.
- [7] M. Menéndez, M. Salicrú, D. Morales, L. Pardo, Divergence measures between populations: Applications in the exponential families, *Communications in Statistic-Theory and Methods*, **26** (1997), 1099-1117.
- [8] O. Parkash, Mukesh, Applications of information measures for the study of Gaussian distribution, *Canadian Journal of Pure and Applied Sciences*, (2013), (Accepted).
- [9] O. Parkash, Mukesh, Development of optimizational principle for minimizing risk in portfolio analysis, *Global and Stochastic Analysis: An International Journal*, **2** (2012), 21-26.
- [10] O. Parkash, Mukesh, New generalized parametric measures of entropy and cross entropy, *American Journal of Mathematics and Sciences*, **1** (2012), 91-96.
- [11] O. Parkash, Mukesh, Relation between information measures and chi-square distribution, *International Journal of Pure and Applied Mathematics*, **84** (2013), 517-524.
- [12] O. Parkash, Mukesh, Two new symmetric divergence measures and information inequalities, *International Journal of Mathematics and Applications*, **4** (2011), 165-179.
- [13] O. Parkash, A.K. Thukral, Statistical measures as measures of diversity, *International Journal of Biomathematics*, **3** (2010), 173-185.

- [14] G.T. Toussaint, Some properties of Matusita's measure of affinity of several distribution, *Annals of the Institute of Statistical Mathematics*, **26** (1974), 389-394.