Mathematics and Statistics: Open Access
Received: Oct 31, 2015, Accepted: Dec 28, 2015, Published: Dec 31, 2015
Math Stat, Volume 1, Issue 1
http://crescopublications.org/pdf/msoa/MSOA-1-002.pdf
Article Number: MSOA-1-002

Research Article Open Access

Chi Square Divergence and Resistor-Average Distance

RamNareshSaraswat* and Shalini Jain

Department of Mathematics and Statistics,Manipal University, Jaipur, Rajasthan -303007, India

***Corresponding Author:** Ram NareshSaraswat, Department of Mathematics and Statistics, Manipal University, Jaipur, Rajasthan -303007, India. E-mail: sarswatrn@gmail.com

Citation: Ram NareshSaraswat and Shalini Jain (2015) Chi Square Divergence and Resistor-Average Distance. Math Stat 1: 002.

Abstract

There are many information and divergence measures are exist in the literature of Information Theory and statistics. These are very useful and play an important role in many areas like as sensor networks, testing the order in a Markov chain, risk for binary experiments, region segmentation and estimation etc. In this paper we shall study to relate chi-square divergence measures to resistor average distance. We have discussed dissimilarity measures between divergence measures.

Keywords: Chi-square divergence measure; New f-divergence measure; Resistor-average distance etc

1. Introduction

Let

$$\Gamma_n = \left\{ P = (p_{1,} p_{2,} \ldots\ldots p_n) \middle| p_i \geq 0, \sum_{i=1}^{n} p_i = 1 \right\}, n \geq 2$$

(1.1)

Be the set of all complete finite discrete probability distributions. There are many information and divergence measures are exist in the literature of Information Theory and Statistics. Csiszar [1, 2] introduced a generalized measure of information using f-divergence measure given by

$$I_f(P,Q) = \sum_{i=1}^{n} q_i f\left(\frac{p_i}{q_i}\right) \quad (1.2)$$

Where $f : \mathbf{R}_+ \to \mathbf{R}_+$ is a convex function and $P, Q \in \Gamma_n$.

As in Csiszar [1, 2], we have interpret undefined expressions by Csiszar's f-divergence is a general class of divergence measures that includes several divergences used in measuring the distance or affinity between two probability distributions. This class is introduced by using a convex function f, defined on (0, ∞). An important property of this divergence is that many known divergences can be obtained from this measure by appropriately defining the convex function f.

2. New f-Divergence Measure and Particular Cases

In this section we shall discuss about some properties of new f-divergence measure for convexity and normalization.

Let

$$\Omega_n = \left\{ P = (p(x)_1, p(x)_{2,} \ldots\ldots p(x)_n) \middle| p(x)_n \geq 0, \int p(x)\, dx = 1 \right\}, n \geq 2$$

be the set of all probability distributions. There are many information and divergence measures exists in the literature on information theory and statistics. In this section we present some properties of a new f-divergence measure [3, 4] and its particular cases which are interesting in areas of information theory and statistics is given by

$$S_f(P,Q) = \int q(x) f\left(\frac{p(x)+q(x)}{2q(x)}\right)(dx) \quad (2.1)$$

Where $f : (1/2, \infty) \to \mathbf{R}_+$ is a convex function and $P, Q \in \Omega_n$.

It is shown that we shall derive some well-known divergence measures using new f-divergence measure such as Chi-square divergence etc. In section 2, we shall discuss new f-divergence measure and their properties. In section 3, we shall discuss chi square divergence measure. In section 4, we can use the Resistor-Average distance as a measure of dissimilarity between two probability densities. In section 5,we have discussed dissimilarity measures between probability density functions

Proposition 2.1 Let $f : [0, \infty) \to \mathbf{R}$ be convex and $P, Q \in \Gamma_n$ then we have the following inequality

$$S_f(P,Q) \geq f(1) \quad (2.2)$$

Equality holds in (2.2) if

$$p(x) = q(x)\ ,\ \forall x \quad (2.3)$$

Proof: If f is a convex function then using well known Jensen's inequality, we get

$$\int [q(x)\, f(X)]\, dX \geq f\left(\int [q(x)\, X)]\, dX\right)$$

Put $X = \dfrac{p(x)+q(x)}{2q(x)}$

$$\int \left[q(x)\, f\left(\frac{p(x)+q(x)}{2q(x)}\right)\right] dx \geq f\left(\int q(x)\, f\left(\frac{p(x)+q(x)}{2q(x)}\right) dx\right)$$

$$\int \left[q(x)\, f\left(\frac{p(x)+q(x)}{2q(x)}\right)\right] dx \geq f\left(\frac{\int p(x)dx + \int q(x)dx}{2\int q(x)dx}\right) \geq 0$$

$$\int \left[q(x)\, f\left(\frac{p(x)+q(x)}{2q(x)}\right)\right] dx \geq f(1)$$

$$\int \left[q(x)\, f\left(\frac{p(x)+q(x)}{2q(x)}\right)\right] dx = S_f(P,Q) \geq f(1)$$

$$S_f(P,Q) \geq f(1)$$

Equality holds if

$$p(x) = q(x) \quad , \forall\ x$$

Corollary 2.1.1 (Non-negativity of new f-divergence measure)

Let $f:[0,\infty)\to\mathbf{R}$

be convex and normalized, i.e.

$$f(1)=0 \qquad (2.4)$$

Then for any $P,Q\in\Gamma_n$ from (2.2) of proposition 2.1 and (2.4),
we have the inequality

$$S_f(P,Q)\geq 0 \qquad (2.5)$$

If f is strictly convex, equality holds in (2.5) if

$$p(x)=q(x)\quad \forall\ x \qquad (2.6)$$

and

$$S_f(P,Q)\geq 0 \text{ and } S_f(P,Q)=0 \text{ if } P=Q \qquad (2.7)$$

Proposition 2.2: Let $f_1\ \&\ f_2$ are two convex functions and
$g=a\,f_1+b\,f_2$ then
$S_g(P,Q)=a\,S_{f_1}(P,Q)+b\,S_{f_2}(P,Q)$
where $a\ \&\ b$ are constants and $P,Q\in\Gamma_n$

Proof: Let $f:[0,\infty)\to\mathbf{R}$ be convex mapping. Let $f_1\ \&\ f_2$ are two convex functions and $g=a\,f_1+b\,f_2$ then we get

$$S_g(P,Q)=\int q(x)g\left(\frac{p(x)+q(x)}{2q(x)}\right)dx=a\left[\int q(x)\,f_1\left(\frac{p(x)+q(x)}{2q(x)}\right)dx\right]+b\left[\int q(x)\,f_2\left(\frac{p(x)+q(x)}{2q(x)}\right)dx\right]$$

$$S_g(P,Q)=\int q(x)g\left(\frac{p(x)+q(x)}{2q(x)}\right)dx=a\,S_{f_1}+b\,S_{f_2}$$

We now give one example of well-known chi square divergence measure which is obtained from new f-divergence measure in the form of continuous probability distributions.

3. Chi Square Divergence Measure

Chi-square divergence measure [5].

If $f(t)=t(t-1)$,
$f'(t)=(2t-1),\ f''(t)=2>0$

Hence function is convex function and normalized $f(1)=0$, then we can apply the properties of new f-divergence measure is given by

$$S_f(P,Q)=\frac{1}{4}\left[\int\left(\frac{p^2(x)}{q(x)}-1\right)dx\right]=\frac{1}{4}\chi^2(P,Q) \qquad (3.1)$$

Where $\chi^2(P,Q)$ is a chi-square divergence measure.

4. Resistor-Average Distance

In this section we shall discussed relation between chi-square divergence measures and resister average distance. This method may be useful in computer science.
Resistor-Average distance as a measure of dissimilarity between two probability densities functions which is defined as

$$D_{RAD}(P,Q)=\left[D_{\chi^2}(P,Q)^{-1}+D_{\chi^2}(Q,P)^{-1}\right]^{-1}$$

It is non-negative and equal to zero if p(x) ≡ q(x), but unlike it, it is symmetric. Another important property of the Resistor-Average distance is that when two classes of patterns S_p and Sq aredistributed according to p(x) and q(x),To see in what manner Resistor-Average distance differs from the Chi-square divergence, it is instructive to consider two special cases: when divergences in both directions between two probability density functions are approximately equal and when one of them is much greater than the other:

$$*D_{\chi^2}(P,Q) \approx D_{\chi^2}(Q,P) \approx D$$

$$D_{RAD}(P,Q) \approx D$$

$$*D_{\chi^2}(P,Q) \approx D_{\chi^2}(Q,P) \approx D$$

$$D_{\chi^2}(P,Q) \approx D_{\chi^2}(Q,P) \ or \ D_{\chi^2}(P,Q) \approx D_{\chi^2}(Q,P)$$

$$D_{RAD}(P,Q) \approx \min D_{\chi^2}(P,Q) \ or \ D_{\chi^2}(Q,P)$$

It can be seen that resistor-average distance very much behaves like a smooth min of $D_{\chi^2}(P,Q)$ and $D_{\chi^2}(P,Q)$).

Actually above method is parallel to results [6-9].

5. Dissimilarity Measures between Probability Density Functions

One of the best known dissimilarity measures between probability density functions (pdfs) is the Chi-square divergence, sometimes also called the Mutual Entropy. It is defined as

$$\chi^2(P,Q) = \frac{1}{4}\int\left[\frac{p^2(x)}{q(x)} - 1\right] dx$$

It is nonnegative and equal to zero if $p(x) = q(x)$. It can be seen that the regions of the image space with a large contribution to the divergence are those in which $p(\mathbf{x})$ is significant and $p(\mathbf{x}) \gg q(\mathbf{x})$. We expect the sets in the training data to be significantly more extensive than the input set, and as a result $p^{(i)}$ to have broader support than $p^{(0)}$. We therefore use $D_{\chi^2}(p^{(0)}, q^{(1)})$ as a "distance measure" between any two data sets.

The appeal of Chi-square divergence stems from its information theory founded approach to quantifying how well a particular probability density function q(x) describes samples from anther probability density function p(x). To gain the intuition behind this divergence, consider from (3.1). The integrand p(x) $\left(\frac{p(x)^2}{q(x)} - 1\right)$ can be seen to have a large value when p(x) is significant and $p(x) \gg q(x)$. Therefore, the regions of the integration space that produce a large contribution to $\chi^2(P,Q)$ be those that are well explained by p(x),but not by q(x) (converse may or may be not true).

This makes $\chi^2(P,Q)$ divergence is may be suitable in cases when it is known a priori that one of the densities p(x) or q(x) describes a wider range of data variation than the other. However, in the recognition framework, this is not the case – pitch and yaw changes are expected to be the dominant modes of variation in any two data.

6. Conclusion

According to this paper we can find out the applications of various information divergence measures in Computer Science and Electronics and Electrical Engineering.

References

[1]. Csiszar I. Information measure,A critical survey.Trans.7th Prague conf. on info. Th. Statist. Decius. Funct, Random Processes and 8th European meeting of statist Volume B. Acadmia Prague, 1978, PP-73-86.

[2]. Csiszar I.Information-type measures of difference of probability functions and indirect observations. studia Sci.Math.hunger.2(1961).299-318.

[3]. Jain KC and RN Saraswat (2012) "A New Information Inequality and its Application in Establishing Relation among various f-Divergence Measures", Journal of Applied Mathematics, Statistics and Informatics (JAMSI), 8(1) (2012), 17-32.

[4]. Jain KC and RN Saraswat (2013) "Some bounds of information divergence measure in term of Relative arithmetic-geometric divergence measure" International Journal of Applied Mathematics and Statistics, 32, (2) (2013), 48-58.

[5]. Pearson K (1900) "On the criterion that a give system of deviations from the probable in the case of correlated system of variables in such that it can be reasonable supposed to have arisen from random sampling", Phil. Mag., 50(1900),157-172.

[6]. Arandjelovi´cOgnjen, Gregory Shakhnarovich, John Fisher, Roberto Cipolla, Trevor Darrell (2005)Face Recognition with Image Sets Using Manifold Density Divergence, Proceedings of the 2005 IEEE Computer Society Conference on Computer Vision and Pattern Recognition (CVPR'05).

[7]. ArandjelovićOgnjen, Roberto Cipolla (2006)An information-theoretic approach to face recognition from face motion manifolds, Image and Vision Computing24(6) 2006, 639–647.
[8]. Barron AR, J. Rissanen and B. Yu (1998)The Minimum Description Length Principle in Coding and Modeling.IEEE Transactions on Information Theory, 44(6):2743–2772, 1998.
[9]. Duda ROP, E Hart and DG Stork.Pattern Classification. John Wiley & Sons, Inc., New York, 2nd edition, 2000.