World Academy of Science, Engineering and Technology
International Journal of Information and Communication Engineering
Vol:5, No:3, 2011

Information Measures Based on Sampling Distributions

Om Parkash, A.K.Thukral, C.P.Gandhi

***Abstract*—**Information theory and Statistics play an important role in Biological Sciences when we use information measures for the study of diversity and equitability. In this communication, we develop the link among the three disciplines and prove that sampling distributions can be used to develop new information measures. Our study will be an interdisciplinary and will find its applications in Biological systems.

***Keywords*—**Entropy; concavity; symmetry; arithmetic mean; diversity; equitability

I. INTRODUCTION

INFORMATION theory has found extensive usage in measurement of diversity of communities. Since the species content and the proportion abundance of species keeps on changing in the communities, diversity indices can be used effectively to study the landscape analysis and development of communities in time and space [1, 2]. There are several measures of diversity [3-6], out of which the most commonly used are due to Shannon [16], Havrada and Charvat [8], Simpson [17] and Renyi [15]. Recently, Parkash and Thukral [7] proved that measures of central tendency and dispersion can be used as information measures. The present work extends the concept developed by the authors to well known sampling distributions.

It is well known that Shannon's [16] mathematical theory of information entropy was introduced to analyze the information carrying capacity of communication channels, serving as a measure of the degree of uncertainty or the extent of ignorance. Information entropy is an extremely important mathematical tool in data compression, signal processing, and communication processes. Entropy is a basic physical quantity that has led to various, and sometimes apparently conflicting, interpretations. It has been successively assimilated to different concepts such as disorder and information. The path-breaking work of well known American Mathematician Shannon [16] who published his first paper "a mathematical theory of communication" is the Magna Carta of the information age. In this paper Shannon [16] introduced the concept of information theoretic entropy by associating uncertainty with every probability distribution $P=\left(p_1, p_2, ..., p_n\right)$ and found that there is a unique function that can measure the uncertainty, is given by

$$H(P) = -\sum_{i=1}^{n} p_i \text{ In } p_i \qquad (1)$$

The probabilistic measure of entropy shown in equation (1) possesses a number of interesting properties. Immediately, after Shannon gave his measure, research workers in many fields saw the potential of the application of this expression and a large number of other measures of information theoretic entropies were derived. Renyi [15] defined entropy of order α as:

$$H_\alpha(P) = \frac{1}{1-\alpha} In\left(\sum_{i=1}^{n} p_i^{\alpha} \Big/ \sum_{i=1}^{n} p_i\right), \alpha \neq 1, \alpha > 0 \qquad (2)$$

which includes Shannon's [16] entropy as a limiting case as $\alpha \to 1$ Zyczkowski [18] explored the relationships between the Shannon's [16] entropy and Renyi's [15] entropies of integer order. The author established a lower and an upper bound for Shannon entropy in terms of Renyi entropies of order 2 and 3.

Havrada and Charvat [8] introduced first non-additive entropy, given by:

$$H^{\alpha}(P) = \frac{\left[\sum_{i=1}^{n} p_i^{\alpha}\right] - 1}{2^{1-\alpha} - 1}, \alpha \neq 1, \alpha > 0 \qquad (3)$$

Lavenda's [9] study is an in-depth analysis of mean entropies, particularly Shannon's [16] and Renyi's [15] entropies, which are expressed as negative logarithms of some means. The functional forms of these entropies follow from the multiplicative law of means. Nanda and Paul [10] derived some ordering results for their own entropy and discussed the properties of the aging classes based on their generalized entropy. The same is shown for Cox's proportional hazard rate model. Rao, Yunmei and Wang [14] used the cumulative distribution of a random variable to define its information content and thereby developed an alternative measure of uncertainty that extends Shannon's [16] entropy to random variables with continuous distributions. Many other developments regarding different measures of entropy have been made by various researchers.

Dr. Om Prakash is working with Department of Mathematics, Guru Nanak Dev University, Amritsar143005,India. E-mail: omparkash777@yahoo.co.in.

Dr. A. K. Thukral is with Department of Botanical & Environmental Sciences, Guru Nanak Dev University, Amritsar 143005 , India, Email: akthukral@rediffmail.com.

C. P. Gandhi is with Department of Mathematics, Guru Nanak Dev University, Amritsar143005, India. Email: cchanderr@gmail.com

II. New Information Measures Based upon Sampling Distributions

A. Measure of entropy based upon χ^2-distribution

We know that χ^2 variate is defined as:

$$\chi^2 = \sum_{i=1}^{n} \frac{\{x_i - M\}^2}{M} \tag{4}$$

Where M is the arithmetic mean of a probability distribution.

The equation (4) can be written as

$$\chi^2 = \sum_{i=1}^{n} \left\{ \frac{x_i - M}{\sum_{i=1}^{n} x_i} \right\}^2 \Bigg/ \frac{M}{\left\{\sum_{i=1}^{n} x_i\right\}^2}$$

$$= \sum_{i=1}^{n} \left\{ \frac{x_i}{\sum_{i=1}^{n} x_i} - \frac{M}{\sum_{i=1}^{n} x_i} \right\}^2 \Bigg/ \frac{\sum_{i=1}^{n} x_i}{n\left\{\sum_{i=1}^{n} x_i\right\}^2}$$

$$= \sum_{i=1}^{n} \left\{ p_i - \frac{1}{n} \right\}^2 \Bigg/ \frac{1}{n\left\{\sum_{i=1}^{n} x_i\right\}} \quad \text{where } p_i = \frac{x_i}{\sum_{i=1}^{n} x_i}$$

$$= n^2 M \sum_{i=1}^{n} \left\{ p_i - \frac{1}{n} \right\}^2$$

$$= n^2 M \left\{ \sum_{i=1}^{n} p_i^2 - \frac{1}{n} \right\}$$

$$\text{or} \quad \frac{\chi^2 + nM}{n^2 M} = \sum_{i=1}^{n} p_i^2 \tag{5}$$

The R.H.S. of equation (5) is a standard measure of information known as measure of energy and has been introduced by Onicescu [11].Thus, we conclude that measure of information can be calculated for known values of χ^2 variate and arithmetic mean and consequently, we have

$$H^1(P) = \frac{\chi^2 + nM}{n^2 M} \tag{6}$$

which is a new measure of information.

B. Measure of entropy based upon t-distribution

We know that t-statistic is defined as

$$t = \frac{|\overline{x} - \mu|}{s/\sqrt{n}} \tag{7}$$

where $\overline{x} = \frac{1}{n}\sum_{i=1}^{n} x_i$ and $s^2 = \frac{1}{n-1}\sum_{i=1}^{n} (x_i - \overline{x})^2$

Now

$$s^2 = \frac{1}{n-1}\left[\sum_{i=1}^{n} x_i^2 - n\overline{x}^2\right]$$

$$= \frac{1}{n-1}\left[\left\{\sum_{i=1}^{n} p_i^2\right\}(n\overline{x})^2 - n\overline{x}^2\right] \text{where } p_i = \frac{x_i}{\sum_{i=1}^{n} x_i}$$

$$= \frac{n\overline{x}^2}{n-1}\left[n\sum_{i=1}^{n} p_i^2 - 1\right], \; n \neq 1 \tag{8}$$

Using (8) in (7), we get

$$\frac{t^2}{n-1} = \frac{(\overline{x} - \mu)^2}{\overline{x}^2\left\{n\sum_{i=1}^{n} p_i^2 - 1\right\}}$$

Taking logarithm both sides, we get

$$\log\frac{t^2}{n-1} = \log\left(\frac{\overline{x} - \mu}{\overline{x}}\right)^2 - \log\left\{n\sum_{i=1}^{n} p_i^2 - 1\right\}$$

$$\text{or} \; -\frac{1}{2}\log\frac{t^2}{n-1} = -\log\left|\frac{\overline{x} - \mu}{\overline{x}}\right| + \log\left\{n\sum_{i=1}^{n} p_i^2 - 1\right\}^{\frac{1}{2}}$$

Hence, we observe that for different values of $n > 2$, the information model for t-distribution becomes

$$I_n(P) = -\frac{1}{2}\log\frac{t^2}{n-1} = -C + \log\left\{n\sum_{i=1}^{n} p_i^2 - 1\right\}^{1/2}, n > 2 \tag{9}$$

where $C = \log\left|\frac{\overline{x} - \mu}{\overline{x}}\right|$.

We choose C such that $\left|\frac{\overline{x} - \mu}{\overline{x}}\right| < 1$ that is,

$$-\overline{x} < \overline{x} - \mu < \overline{x} \quad or \; \overline{x} > \frac{\mu}{2} \; and \; \mu > 0$$

Without any loss of generality, we can assume that

$$C < \left\{n\sum_{i=1}^{n} p_i^2 - 1\right\}^{1/2}$$

Now, we show that the R.H.S. of equation (9) is a theoretical measure of information. For this purpose, we have studied the following properties:

(i) Obviously $I_n(P) \geq 0$ as $C<0$ and

$$\left\{n\sum_{i=1}^{n} p_i^2 - 1\right\}^{1/2} < 1 \ \forall p_i$$

(ii) $I_n(P)$ is a continuous function of p_i.

(iii) $I_n(P)$ is permutationally symmetric function of p_i.

(iv) *Concavity*: To study its concavity, we proceed as follows:

We have $\dfrac{\partial I_n(P)}{\partial p_i} = \dfrac{np_i}{n\sum_{i=1}^{n} p_i^2 - 1}$

Also $\dfrac{\partial^2 I_n(P)}{\partial p_i^2} = \dfrac{n}{\left\{n\sum_{i=1}^{n} p_i^2 - 1\right\}^2}\left[\mathrm{n}\sum_{i=1}^{n} p_i^2 - 2np_i^2 - 1\right]$

It has been verified numerically that

$$\left[\mathrm{n}\sum_{i=1}^{n} p_i^2 - 2np_i^2 - 1\right] < 0 \quad \forall \mathrm{n}, \mathrm{p_i}$$

Thus, we observe that $\dfrac{\partial^2 I_n(P)}{\partial p_i^2} < 0$

which shows that $I_n(P)$ is a concave function of p_i .

(v) For obtaining maximum value, we consider the Lagrange function given by

$$g(P) = -\frac{1}{2}\log\frac{t^2}{n-1} = -C + \log\left\{n\sum_{i=1}^{n} p_i^2 - 1\right\}^{1/2} - \lambda\left(\sum_{i=1}^{n} p_i - 1\right)$$

Thus, $\dfrac{\partial g(P)}{\partial p_1} = \dfrac{np_1}{n\sum_{i=1}^{n} p_i^2 - 1} - \lambda$

$$\frac{\partial g(P)}{\partial p_2} = \frac{np_2}{n\sum_{i=1}^{n} p_i^2 - 1} - \lambda$$

.

.

$$\frac{\partial g(P)}{\partial p_n} = \frac{np_n}{n\sum_{i=1}^{n} p_i^2 - 1} - \lambda$$

For maximum value, putting

$\dfrac{\partial g(P)}{\partial p_1} = \dfrac{\partial g(P)}{\partial p_2} = \dfrac{\partial g(P)}{\partial p_3} = \ldots \dfrac{\partial g(P)}{\partial p_n} = 0$, which gives

$$\frac{np_1}{n\sum_{i=1}^{n} p_i^2 - 1} = \frac{np_2}{n\sum_{i=1}^{n} p_i^2 - 1} = \ldots = \frac{np_n}{n\sum_{i=1}^{n} p_i^2 - 1}$$

which further gives $p_1 = p_2 = \ldots\ldots = p_n$.

Also using $\sum_{i=1}^{n} p_i = 1$, we get $p_i = \dfrac{1}{n}$. Thus the maximum value arises when the distribution is uniform.

Under the above properties, we see that $I_n(P)$ will be an information measure and consequently, we conclude that if t is any t-statistic, then $-\dfrac{1}{2}\log\dfrac{t^2}{n-1}$ will act as an information measure.

C. Measure of entropy based upon F-distribution.

The F-statistic is given by $F = \dfrac{s_x^{\,2}}{s_y^{\,2}}, s_x^{\,2} > s_y^{\,2}$

where $s_x^{\,2} = \dfrac{1}{n_1 - 1}\sum_{i=1}^{n_1}\left(x_i - \overline{x}\right)^2$ and

$$s_y^{\,2} = \frac{1}{n_2 - 1}\sum_{j=1}^{n_2}\left(y_j - \overline{y}\right)^2 \ , \ n_1 > 2, n_2 > 2$$

Now $s_x^{\,2} = \dfrac{1}{n_1 - 1}\left[\sum_{i=1}^{n_1} x_i^2 - n_1\overline{x}^2\right]$

$$= \frac{1}{n_1 - 1}\left[\left\{\sum_{i=1}^{n_1} p_i^2\right\}(n_1\overline{x})^2 - n_1\overline{x}^2\right],$$

where $p_i = \dfrac{x_i}{\sum_{i=1}^{n} x_i}$

$$= \frac{n_1\overline{x}^2}{n_1 - 1}\left[n_1\sum_{i=1}^{n_1} p_i^2 - 1\right]$$

Similarly, we have

$$s_y^{\,2} = \frac{n_2\overline{x}^2}{n_2 - 1}\left[n_2\sum_{j=1}^{n_2} p_j^2 - 1\right]$$

Thus,

$$F = \frac{s_x^{\,2}}{s_y^{\,2}} = \frac{n_1\left(n_2 - 1\right)}{n_2\left(n_1 - 1\right)}\frac{\overline{x}^2}{\overline{y}^2}\frac{n_1\sum_{i=1}^{n_1} p_i^{\,2} - 1}{n_2\sum_{j=1}^{n_2} p_j^{\,2} - 1} \ , \ n_2 > n_1$$

Taking logarithm both sides, we get

$$\log F = \log \frac{n_1(n_2-1)}{n_2(n_1-1)}\frac{\bar{x}^2}{\bar{y}^2} + \log \frac{n_1\sum_{i=1}^{n_1} p_i^2 - 1}{n_2\sum_{j=1}^{n_2} p_j^2 - 1}$$

or $\log F = C + \log A$

where $C = \log \frac{n_1(n_2-1)}{n_2(n_1-1)}\frac{\bar{x}^2}{\bar{y}^2}$, and $A = \frac{n_1\sum_{i=1}^{n_1} p_i^2 - 1}{n_2\sum_{j=1}^{n_2} p_j^2 - 1}$

We, now propose a new measure of information based on F-distribution, given by

$$_n\psi(P) = C + \log \frac{n_1\sum_{i=1}^{n_1} p_i^2 - 1}{n_2\sum_{j=1}^{n_2} p_j^2 - 1}, n_2 > n_1 \qquad (10)$$

Now, we show that the R.H.S. of equation (10) is a theoretical measure of information. For this purpose, we study the following properties:

(i) Obviously $_n\psi(P) \geq 0$ as $C > 0$,

$$\left\{n_1\sum_{i=1}^{n_1} p_i^2 - 1\right\}^{1/2} < 1 \ \forall p_i \text{ and}$$

$$\left\{n_2\sum_{j=1}^{n_2} p_j^2 - 1\right\}^{1/2} < 1 \ \forall p_j$$

(ii) $_n\psi(P)$ is a continuous function of p_i and p_j.

(iii) $_n\psi(P)$ is a symmetric function of p_i and p_j.

(i) **Concavity:** To study its concavity, we have

$$_n\psi(P) = C + \log\left(n_1\sum_{i=1}^{n_1} p_i^2 - 1\right) - \log\left(n_2\sum_{j=1}^{n_2} p_j^2 - 1\right)$$

Take $_n\psi^1(P) = \log\left(n_1\sum_{i=1}^{n_1} p_i^2 - 1\right)$ and

$$_n\psi^2(P) = \log\left(n_2\sum_{j=1}^{n_2} p_j^2 - 1\right)$$

Thus

$$\frac{\partial\, _n\psi^1(P)}{\partial p_i} = \frac{2n_1 p_i}{n_1\sum_{i=1}^{n_1} p_i^2 - 1}$$

Also,

$$\frac{\partial^2\, _n\psi^1(P)}{\partial p_i^2} = \frac{2n_1}{\left\{n_1\sum_{i=1}^{n_1} p_i^2 - 1\right\}^2}\left[n_1\sum_{i=1}^{n_1} p_i^2 - 2n_1 p_i^2 - 1\right] < 0 \forall n_1 > 2$$

Thus, $_n\psi^1(P)$ is a concave function of p_i

Similarly, $$\frac{\partial^2\, _n\psi^2(P)}{\partial p_j^2} = \frac{2n_2}{\left\{n_2\sum_{j=1}^{n_2} p_j^2 - 1\right\}^2}\left[\mathrm{n}_2\sum_{j=1}^{n_2} p_j^2 - 2n_2 p_j^2 - 1\right]$$

$$< 0 \forall n_2 > 2$$

Hence, $_n\psi^2(P)$ is a concave function of p_j. As difference of two concave functions is also a concave function, we conclude that $_n\psi(P)$ is a concave function.

(v) For obtaining maximum value,we consider the Lagrange function:

$$f(p) = C + \log\left\{n_1\sum_{i=1}^{n_1} p_i^2 - 1\right\} - \log\left\{n_2\sum_{j=1}^{n_2} p_j^2 - 1\right\} - \lambda\left(\sum_{i=1}^{n_1} p_i - 1\right) - \mu\left(\sum_{j=1}^{n_2} p_j - 1\right)$$

Thus, we have

$$\frac{\partial f(p)}{\partial p_1} = \frac{2n_1 p_1}{n_1\sum_{i=1}^{n_1} p_i^2 - 1} - \frac{2n_2 p_1}{n_2\sum_{j=1}^{n_2} p_j^2 - 1} - \lambda - \mu$$

$$\frac{\partial f(p)}{\partial p_2} = \frac{2n_1 p_2}{n_1\sum_{i=1}^{n_1} p_i^2 - 1} - \frac{2n_2 p_2}{n_2\sum_{j=1}^{n_2} p_j^2 - 1} - \lambda - \mu$$

.
.
.

$$\frac{\partial f(p)}{\partial p_n} = \frac{2n_1 p_n}{n_1\sum_{i=1}^{n_1} p_i^2 - 1} - \frac{2n_2 p_n}{n_2\sum_{j=1}^{n_2} p_j^2 - 1} - \lambda - \mu$$

For maximum value, putting

$\frac{\partial f(p)}{\partial p_1} = \frac{\partial f(p)}{\partial p_2} = \frac{\partial f(p)}{\partial p_3} = \ldots \frac{\partial f(p)}{\partial p_n} = 0$, we get

$$2p_1\left(\frac{n_1}{n_1\sum_{i=1}^{n_1} p_i^2 - 1} - \frac{n_2}{n_2\sum_{j=1}^{n_2} p_j^2 - 1}\right)$$

$$= 2p_2\left(\frac{n_1}{n_1\sum_{i=1}^{n_1} p_i^2 - 1} - \frac{n_2}{n_2\sum_{j=1}^{n_2} p_j^2 - 1}\right)$$

$$\ldots = 2p_n\left(\frac{n_1}{n_1\sum_{i=1}^{n_1} p_i^2 - 1} - \frac{n_2}{n_2\sum_{j=1}^{n_2} p_j^2 - 1}\right)$$

which gives $p_1 = p_2 = \ldots = p_n$.

Also using $\sum_{i=1}^{n_1} p_i = 1, \sum_{j=1}^{n_2} p_j = 1$, we get

$$p_i = \frac{1}{n_1} \ and \ p_j = \frac{1}{n_2}.$$

Thus, we see that the maximum value arises when the distribution is uniform.

Under the above properties, we see that ${}_n\psi(P)$ introduced above is an information theoretic measure and consequently, we conclude that if F is any F-statistic, then will act as an information measure.

III. Discussion

The literature of information theory deals with the development of information theoretic measures- probabilistic as well fuzzy. We observe that there has been redundancy and overlapping in similar situations, which, if removed, can increase the efficiency of the system. The development of new probabilistic and fuzzy measures will definitely reduce uncertainty, which will help to increase the efficiency and remove uncertainty for the betterment of mankind. Some of the fuzzy measures have recently been developed by Parkash, Sharma and Mahajan [12, 13] and have successfully been applied for the study of maximum entropy principles. But, there are a variety of Mathematical, Statistical and Biological disciplines where we need a variety of information measures to extend the scope of their applications. Keeping this idea in mind, we have developed some measures of information and concluded that for the known values of χ^2 variate, arithmetic mean, t -statistic and F-statistic, the information content of a discrete frequency distribution can be calculated and consequently, new information measures can be developed.

Acknowledgements

The senior author is thankful to University Grants Commission, New Delhi, for providing the financial assistance.

References

[1] K. McGarigal and B.J. Marks, 1995. FRAGSTATS: Spatial Pattern Analysis Program for Quantifying Landscape Structure. General Technical Report (USDA Forestry Service, Corvallis, 1995).
[2] A. Chawla, S. Rajkumar, K.N. Singh, Brij Lal, R.D. Singh and A.K. Thukral, Plant species diversity along an altitudinal gradient of Bhabha valley in western Himalaya. Journal of Mountain Science 5 (2008) 157-177.
[3] A.E. Magurran, Ecological Diversity and its Measurement (Croom Helm, 1988, London).
[4] M, Gadgil and Meher-Homji, V.M. Ecological diversity. In: J.C. Daniel and J.S. Serrao (Eds.) Proceedings of the Centenary Seminar of the Bombay Natural History Society (Oxford University Press, Bombay, 1990).
[5] S.Baumgartner, Measuring Biodiversity (University of Heidelberg, Heidelberg, 2005).
[6] P.D.S. Kinako, Mathematical elegance and ecological naivety of diversity indices, African Journal of Ecology 21 (2009) 93-99.
[7] O.Parkash and A.K.Thukral, Statistical measures as measures of diversity, International Journal of Biomathematics 2009 (In press).
[8] J.H. Havrada and F.Charvat, Quantification methods of classification process:Concept of structural □-entropy, Kybernetika 3, (1967) 30-35.
[9] B. H. Lavenda, Mean Entropies, Open System Infor. Dyn. 12, (2005) 289-302.
[10] A.K. Nanda. and P. Paul, Some results on generalized residual entropy, Information Sciences 176 (2006) 27-47.
[11] O. Onicescu, Energie Informationelle (C.R.Academic Science, 1966).
[12] O. Parkash, P. K. Sharma and R. Mahajan, New measures of weighted fuzzy entropy and their applications for the study of maximum weighted fuzzy entropy principle, Information Sciences 178 (2008) 2389-2395.
[13] O. Parkash, P. K. Sharma and R. Mahajan, Optimization principle for weighted fuzzy entropy using unequal constraints, Southeast Asian Bulletin Mathematics 2008 (In Press).
[14] M.C. Rao, V.B.C.,Yunmei and F.Wang, Commulative residual entropy: a new measure of Information, IEEE Trans. Inform. Theory 50 (2004) 1220-1228.
[15] A. Renyi , On measures of entropy and information, Proc. 4th Ber. Symp. Math. Stat. and Prob. 1 (1961) 547-561.
[16] C. E. Shannon, A mathematical theory of communication, Bell. Sys. Tech. Jr. 27(1948) 379-423, 623-659.
[17] E. H. Simpson, Measurement of diversity, Nature 163 (1949) 688.
[18] K. Zyczkowski, Renyi extrapolation of Shannon entropy, Open Syst. Inf. Dyn., 10 (1948) 297-310.