Volume-04
Issue-03
March -2019

ISSN: 2455-3085 (Online)
RESEARCH REVIEW International Journal of Multidisciplinary
www.rrjournals.com [UGC Listed Journal]

Measuring Information Content through Measures of Central Tendency

[1]Vishal & [2]Prachi

[1,2]Department of Mathematics, S.S.M. College, Dinanagar (India)

ARTICLE DETAILS

Article History
Published Online: 13 March 2019

Keywords
Entropy; concavity; symmetry; arithmetic mean; geometric mean; harmonic mean; equitability

ABSTRACT

It is well known that Statistics is the field which is extensively used for the measurement of central tendency, dispersion, comparison and covariation. On the other hand, measures of information are used to study diversity and equitability. These two fields have been used independent of each other for data analysis. In this communication, we develop the link between the two and prove that the well known standard statistical measures can be used as information measures. Our study will be a new interdisciplinary field of research and it will be possible to describe information content of a system from its statistics.

1. INTRODUCTION

The most successful measure of information, which is interpreted as the uncertainty content of a random experiment is due to Shannon's [6]. In fact, one can approach Shannon's [6] measure of entropy with two different approaches. The first approach, which is preferred by coding theorists, is the approach under which, Shannon entropy is the solution to a coding-theoretic problem, and represents the ideal rate of compression. The second one is the axiomatic approach which is more relevant to the point of view of information measures in the strict sense. In this case, one lists axioms, which an adequate information measure should possess. It was an Shannon who in 1948, remarked that since uncertainty is always associated with every probability

distribution $P = (p_1, p_2, ..., p_n)$, the uncertainty measure must be a function of probabilities and with certain desirable postulates, investigated and found that the only function which satisfies these postulates is given by

$$H(P) = -\sum_{i=1}^{n} p_i \operatorname{In} p_i \tag{1.1}$$

He called the measure (1.1) as entropy and studied its many interesting properties which, later on proved to be extremely useful in many disciplines of Mathematical Sciences. It was Schroeder [5], who pointed out that Shannon entropy, while useful in communications theory, is conceptually inadequate as a measure of information. After the invention of Shannon's [6] measure of entropy, many scientists became interested in the field of information theory. Consequently, Renyi [4] introduced entropy of order α, given by:

$$H_a(P) = \frac{1}{1-\alpha} \operatorname{In}\left[\sum_{i=1}^{n} p_i^a \Big/ \sum_{i=1}^{n} p_i\right], a \neq 1, a > 0 \tag{1.2}$$

Havrada and Charvat [2] introduced another non-additive measure of entropy, given by:

$$H^{\alpha}(P) = \frac{1}{1-\alpha}\left[\sum_{i=1}^{n} p_i^{\alpha}\right], \alpha \neq 1, \alpha > 0 \tag{1.3}$$

Burg [1] developed his non-parametric measure of entropy, given by

$$H^1(P) = \sum_{i=1}^{n} \log p_i \tag{1.4}$$

There exist many well-known measures of information which are frequently used by the researchers working in Biological Sciences for measuring diversity and equitability of different communities. Some of these measures are due to Shannon [6], Renyi [4], Simpson [7], Weiner [8] etc. Of course Shannon's [6] measure is most widely applicable and possesses many interesting and desirable properties. But, it has a limitation that it deals with exponential families only whereas in actual practice, there are many distributions which are non-exponential, there are many growth curves which do not follow exponential law. Thus, there is a need for developing new measures to extend the scope of their applications. Recently, Parkash and Thukral [3] have developed such measures which find tremendous applications and are definitely helpful to the researchers in the field of Biology. We have taken the motivation from this work and extended their idea for the development of this paper.

2. NEW INFORMATION MEASURES BASED UPON CENTRAL TENDENCY

In this section, we introduce some new probabilistic measures of information depending upon geometric mean, arithmetic mean and harmonic mean of a discrete distribution.

I. Information Measure in terms of Measures of Central Tendency

Let a random variable X takes values $x_1, x_2, x_3, \ldots, x_n$. Then geometric mean G, arithmetic mean M and harmonic mean H of these n observations are given by:

$$G = (x_1 . x_2 . x_3 ...... x_n)^{\frac{1}{n}} \qquad x_i \geq 0 \tag{2.1}$$

$$M = \frac{1}{n} \sum_{i=1}^{n} x_i \tag{2.2}$$

$$H = \frac{n}{\frac{1}{x_1} + \frac{1}{x_2} + \ldots + \frac{1}{x_n}} = \frac{n}{\sum_{i=1}^{n} \frac{1}{x_i}} \tag{2.3}$$

Equations (2.1), (2.2) and (2.3) can be rewritten as

$$G = (p_1 . p_2 . p_3 ...... p_n)^{\frac{1}{n}} \sum_{i=1}^{n} x_i \tag{2.4}$$

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

$$nM = \sum_{i=1}^{n} x_i \tag{2.5}$$

$$\frac{H}{M} = \frac{n^2}{\sum_{i=1}^{n} \frac{1}{p_i}} \tag{2.6}$$

Dividing equations (2.6) and (2.4), we get

$$\frac{H}{MG}=\frac{n}{\left\{\sum_{i=1}^{n} x_i\right\}(p_1 \cdot p_2 \cdot p_3 \ldots p_n)^{\frac{1}{n}} \sum_{i=1}^{n} \frac{1}{p_i}} \tag{2.7}$$

Equation (2.7) further reduces to

$$\frac{H}{nG}=\frac{1}{(p_1 \cdot p_2 \cdot p_3 \ldots p_n)^{n} \sum_{i=1}^{n} \frac{1}{p_i}} \tag{2.8}$$

Taking logarithms, we get

$$\log\left(\frac{H}{nG}\right)=-\frac{1}{4}\sum_{i=1}^{n}\log p_i \cdot \log\left(\sum_{i=1}^{n}\frac{1}{pi}\right) \tag{2.9}$$

which gives the relation between harmonic mean H and geometric mean G in terms of probabilities. We shall prove that this relation gives a new measure of information.

Now, dividing equations (2.1) and (2.2), we get

$$\frac{M}{G}=\frac{\left\{\sum_{i=1}^{n} x_i\right\}}{n(p_1 \cdot p_2 \cdot p_3 p_n)^{1/4}\left\{\sum_{i=1} x_i\right\}} \tag{2.10}$$

Equation (2.10) upon further simplification reduces to

$$-n\log\left(\frac{nM}{G}\right)=\sum_{i=1}^{n}\log\ p_i \tag{2.11}$$

which gives the relation between arithmetic mean M and geometric mean G in terms of probabilities. Since the result introduced in the R.H.S. of equation (2.11) corresponds to Burg's [1] standard measure of entropy, we therefore conclude that for the given values of arithmetic mean and geometric mean, the information content of the probability distribution can be obtained.

Again, subtracting equations (2.9) and (2.11), we get

$$\log\left[\frac{H}{n^2M}\right]=-\log\sum_{i=1}^{n}\frac{1}{p_i} \tag{2.12}$$

which gives the relation between harmonic mean H and arithmetic mean M in terms of probabilities. We shall prove that this relation further gives a new measure of information.

Thus, we consider the following functions and show that these represent information measures:

(I) $$\zeta_n(P)=-\log\left(\sum_{i=1}^{n}\frac{1}{p_i}\right) \qquad n\ge 2,\ 0<p_i\le 1 \tag{2.13}$$

(II) $$\eta_n(P) = -\frac{1}{n}\sum_{i=1}^{n}\log p_i - \log\left[\sum_{i=1}^{n}\frac{1}{p_i}\right], n \geq 2, 0 < p_i \leq 1 \qquad (2.14)$$

To prove its authenticity of these measures, we proceed as follows:

(i) $\xi_n(P)$ is permutationaly symmetric as it does not changes if $p_1, p_2, p_3, \ldots, p_n$ are rearranged among themselves. This property is desirable since the labeling of the outcomes should not affect the entropy.

(ii) Since $\frac{1}{p_i}$ is continuous function of $p_1, p_2, p_3, \ldots, p_n$ for $0 < p_i < 1$, $\xi_n(P)$ is also continuous everywhere in the same interval.

(iii) Clearly $\xi_n(P) < 0$

(iv) We have $$\xi''(P) = -\left[\frac{1}{P_i^4\left[\sum_{i=1}^{n}\frac{1}{p_i}\right]}\left\{2p_i\left[\sum_{i=1}^{n}\frac{1}{p_i}\right]-1\right\}\right]$$

Now, since $p_i.\left(\sum_{i=1}^{n}\frac{1}{p_i}\right) - 1 > 0$ always, we see that $\xi_n(P) < 0$

Thus, $\xi_n(P)$ is a concave function of $p_1, p_2, p_3, \ldots, p_n$. This is very useful property since a local maximum will also be the global maximum for a concave function.

(v) **For maximum value**, we consider the following Lagrangian:

$$L = -\log\left[\sum_{i=1}^{n}\frac{1}{p_i}\right] - \lambda\left[\sum_{i=1}^{n}p_i - 1\right]$$

Thus ,we have $\frac{\partial L}{\partial p_i} = \frac{1}{\sum_{i=1}^{n}\frac{1}{p_i}}\frac{1}{p_1^2} - \lambda$, $\frac{\partial L}{\partial p_2} = \frac{1}{\sum_{i=1}^{n}\frac{1}{p_i}}\frac{1}{p_2^2} - \lambda, \ldots, \frac{\partial L}{\partial p_i} = \frac{1}{\sum_{i=1}^{n}\frac{1}{p_i}}\frac{1}{p_i^2} - \lambda$

Now, $\frac{\partial L}{\partial p_i} = 0$ gives that $\frac{\partial L}{\partial p_1} = \frac{\partial L}{\partial p_2} \ldots = \frac{\partial L}{\partial p_1}$

This is possible if and only if $p_1 = p_2 = p_3 = \ldots = p_n$

Also $\sum_{i=1}^{n} p_i = 1 \Rightarrow n.p_i = 1 \Rightarrow p_i = \frac{1}{n}$

Thus, the maximum value of $\xi_n(P)$ will exist at $p_i = \frac{1}{n}$ and is given by

$$\left[\xi_n(P)\right]_{max} = -\log\left\{\sum_{i=1}^{n} n\right\} = -\log n^2 = -2\log n$$

Hence, we see that introduced in equation (2.13) satisfies all the essential properties of an information measure, it is a new measure of information. It is thus concluded that for the given values of arithmetic mean and harmonic mean, the information content of the probability distribution can be obtained. Next, with the help of the data, we have presented the measure (2.13) graphically. For this purpose, we have fixed n =2, then for different probabilities, we have computed different values of $\xi_n(P)$ as show in the table-2.1.

Table-2.1

p_1	p_2	$\xi_n(P)$
0.05	0.95	-4.3959
0.15	0.85	-2.9714
0.30	0.70	-2.2515
0.40	0.60	-2.0589
0.50	0.50	-2.0000
0.60	0.40	-2.0589
0.70	0.30	-2.2515
0.80	0.20	-2.6439
0.90	0.10	-3.4737
0.95	0.05	-4.3959

Consequently, we have obtained the following Fig.-2.1 which shows that the measure introduced in equation (2.13) is a concave function.

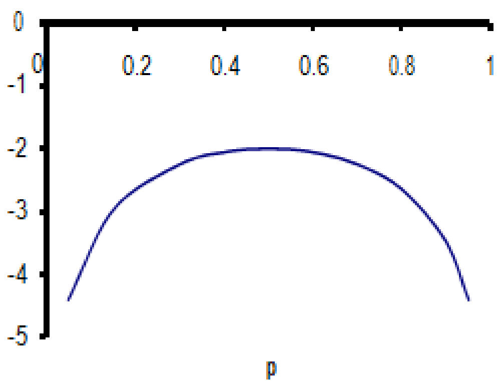

Fig-2.1

Next, to prove the essential and desirable properties of another measure introduced in (2.14), we proceed as follows:

(i) $\eta_n(P)$ is permutationaly symmetric.

(ii) Since $\frac{1}{p_i}$ is continuous function of $p_1, p_2, p_3, \ldots, p_n$ for $0 < p_i < 1$, $\eta_n(P)$ is also continuous everywhere in the same interval.

(iii) Obviously $\eta_n(P) < 0$

Since the measure introduced in (2.14) corresponds to Burg's [] measure of entropy, which gives negative value, the measure of information introduced in (2.14) shall be negative.

(iv) We have

(v)

$$\eta_{n(P)} = -\frac{1}{p_i^2}\left[-\frac{1}{n} + \frac{1}{p_i^2\left(\sum_{i=1}^{n}\frac{1}{p_i}\right)}\left\{2p_i\left(\sum_{i=1}^{n}\frac{1}{p_i}\right) - 1\right\}\right]$$

Now, since $p_i \sum_{i=1}^{n} \frac{1}{p^i} - 1 > 0$ always, and also it has numerically been verified that the second term within brackets is always greater than the first term, we see that $\eta_n''(P) < 0$ Thus, $\eta_n(P)$ is a concave function of $p_1, p_2, p_3, \ldots, p_n$.

(vi) **For maximum value**, we consider the following Lagrangian:

$$L = -\frac{1}{4}\sum_{i=1}^{n}\log p_i - \log\left(\sum_{i=1}^{n}\frac{1}{p_i}\right) - \lambda\left(\sum_{i=1}^{n} p_i - 1\right)$$

As proved above, we see that the maximum value of $\eta_n(P)$ will exist at $p_i = 1/n$ and is given by

$$\eta_n[(P)]_{max} = -\frac{1}{n}\sum_{i=1}^{n}\log\left(\frac{1}{n}\right) - \log\left(\sum_{i=1}^{n} n\right) = \log n - 2\log n = -\log n$$

Hence, we see that $\eta_n(P)$ introduced in equation (2.14) satisfies all the essential properties of an information measure, it is a new measure of information. It is thus concluded that for the given values of arithmetic mean and harmonic mean, the information content of the probability distribution can be obtained. Next, with the help of the data, we have presented the measure (2.14) graphically and obtained the following Fig.-2.2 which shows that the measure introduced in equation (2.14) is a concave function.

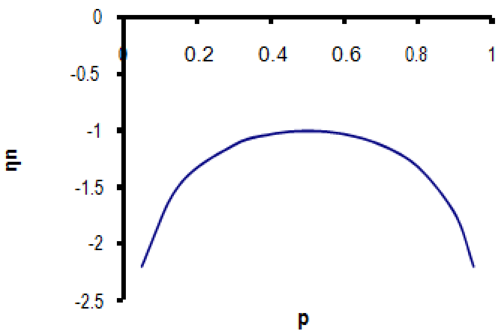

Fig-2.2

REFERENCES

1. Burg, J.P. (1972). The relationship between maximum entropy spectra and maximum likelihood spectra. In: Childrers, D.G.(ed). Modern Spectral Analysis. Pp.130-131.
2. Havrada, J.H. and Charvat, F. (1967). Quantification methods of classification process: Concept of structural □-entropy. Kybernetika **3**: 30-35.
3. Parkash, O. and Thukral, A.K. (2010). Statistical measures as measures of diversity. International Jr. of Biomathematics, **3**(2), 173-185.
4. Renyi, A. (1961). On measures of entropy and information. Proceedings 4th Berkeley Symposium on Mathematical Statistics and Probability **1**: 547-561.
5. Schroeder, M. J. (2004). An alternative to entropy in the measurement of information. Entropy **6**: 388-412.
6. Shannon, C. E.(1948). A mathematical theory of communication. Bell System Technical Journal **27**: 379-423, 623-659.
7. Simpson, E.H. (1949). Measurement of diversity. Nature **163**: 688.
8. Wiener, N. (1949). The extrapolation, interpolation and smoothing of stationary time series with engineering applications. John Wiley and Sons, New York.