International Journal of Science and Research (IJSR)
ISSN: 2319-7064
SJIF (2022): 7.942

P-Norm and Some of its Basic Properties: A Review

Neha Bhardwaj

Assistant Professor at Arya (P.G.) College Panipat, Haryana, India
Email: *bhardwajneha15[at]gmail.com*

Abstract: ***A number of information measures, including Shannon's information measure and the measures of order γ and type δ, have been created and thoroughly examined. These metrics are used in coding theory, statistics, and pattern recognition, among other fields. This research aims to examine the fundamental characteristics of the P-Norm information measure, which is a generalized class of information measures as noted by Arimoto. Many interesting algebraic and analytic features of this measure will also be discussed. It is also explored whether Shannon's measure may be derived from the P-norm information measure as the limiting case for P → 1. It is also discussed that the measure lacks the additivity feature, even if it displays a form of additivity that we will call Pseudo-additivity.***

Keywords: Shannon's information measure, P-Norm information measure, coding theory, pattern recognition, Pseudo-additivity

1. Introduction

Numerous information measures, including Shannon's information measure (C. E. Shannon 1948) and the measures of order γ and type δ presented by Alfred Renyi (1961), Havrda and Charvát (1967), and Daróczy (1970), have been produced and intensively investigated. These metrics are used in coding theory, statistics, and pattern recognition, among other fields.

This work aims to examine the fundamental characteristics of the P-Norm information measure, which was identified by Arimoto (1971) as an instance of a broader class of information measures. Initially, we provide an axiomatic description of this measurement. Additionally, a variety of fascinating analytic and algebraic aspects of this measure will be covered.

The ability to derive Shannon's measure as the limiting case for P → 1 from the P-norm information measure is a crucial feature. Additionally, we demonstrate that the measure exhibits a sort of additivity that we will refer to as Pseudo-additivity rather than the additivity condition.

2. Definition

For convenience, we start by introducing a few notations.
The set of Positive real numbers that are not equal to one will be referred to frequently. This set is indicated by ℝ with

$$\mathbb{R} = \{P : P > 0, P \neq 1\}$$

Additionally, we define Δ_n as all of the nary Probability distributions together
$L = (l_1, l_2, l_3, \dots, l_n)$ which meet the requirements

$$l_\alpha \geq 0 \,;\ \alpha = 1,2,3, \dots \dots, n$$
$$\sum_{\alpha=1}^{n} l_\alpha = 1$$

Definition 1: The distribution's P-norm information L has a definition for
$P \in \mathbb{R}$ by

$$H_P(L) = \frac{P}{P-1}\left[1 - \left[\sum_{\alpha=1}^{n} {l_\alpha}^P\right]^{\frac{1}{P}}\right] \qquad (1)$$

A real function $\Delta_n \to \mathbb{R}$ defined on Δ_n, where $n \geq 2$ and ℝ is the set of real numbers, is the P-norm information measure. We can inquire as to what characteristics the function must have in order to provide an axiomatic characterization.

$I : \Delta_n \to \mathbb{R}$ such that $I(L) = H_P(L)$ for all $L = (l_1, l_2, l_3, \dots, l_n) \in \Delta_n$ and $P \in \mathbb{R}$.

Theorem 1. Let P ϵ ℝ be a real function and let I: $\Delta_n \to \mathbb{R}$. & Let I permit to meet these postulates.

a) $I(L) = a[\sum_{\alpha=1}^{n} F(l_\alpha)]^{\frac{1}{P}} + b;\ a, b \neq 0.$
b) $F(.)$ *is a continous function on* $[0,1)$.
c) *If* $L \in \Delta_n$ *and* $M \in \Delta_m$ *are stochastically independent then I is pseudo − additive*:

$$I(LM) = I(L) + I(M) - \frac{1}{b} I(L).I(M)$$

d) $I\left(\frac{1}{n}, \frac{1}{n}, \frac{1}{n}, \dots \dots \frac{1}{n}\right) = \frac{p}{P-1}\left(1 - n^{\frac{1-P}{P}}\right)$

where arbitrary constants a and b are used. For every L that lies inside Δ_n, $I(L) = H_P(L)$.

Proof: For $L \in \Delta_n$ and $M \in \Delta_m$, it follows from combining Postulates (i) and (iii).

$$\left[\sum_{\alpha=1}^{n}\sum_{\beta=1}^{m} F(l_\alpha m_\beta)\right]^{\frac{1}{P}} = -\frac{a}{b}\left[\sum_{\alpha=1}^{n} F(l_\alpha)\right]^{\frac{1}{P}} . \left[\sum_{\beta=1}^{m} F(m_\beta)\right]^{\frac{1}{P}} \qquad (2)$$

Next, we'll demonstrate that (2)'s answer is provided by

$$F(x) = \left(-\frac{a}{b}\right)^P . x^k \qquad (3)$$

Let c, d, e and h be integers such that $1 \leq e \leq c$ and $1 \leq h \leq d$. If we substitute $n = c - e + 1$, $m = d - h + 1$, $L_\alpha = 1/c$, $\alpha = 1, 2 \dots c - e$, $L_{c-e+1} = e/c$, $q_\beta = 1/d$, $\beta = 1,2, \dots, d - h$ and $q_{d-h+1} = h/d$

into (**Error! Reference source not found.**) we put ν (x) = xF(1/x) and multiplying both sides of (**Error! Reference source not found.**) with cs ,we get

$$(c-d)(d-h)\nu(cd)+h(c-e)\nu\left(\frac{cd}{h}\right)+e(d-h)\nu\left(\frac{cd}{e}\right)+eh\nu\left(\frac{cd}{eh}\right) = \left(-\frac{a}{b}\right)^P\left[(c-e)\nu(c)+e\nu\left(\frac{c}{e}\right)\right].\left[(d-h)\nu(d)+h\nu\left(\frac{d}{h}\right)\right] \quad (4)$$

Setting e= h = 1 in (**Error! Reference source not found.**) yields

$$\nu(cd)=\left(-\frac{a}{b}\right)^P \nu(c)\nu(d) \quad (5)$$

Next, by using (5) and substituting e = 1 resp. h = 1 into equation (4), we can derive

$$\nu\left(\frac{cd}{h}\right)=\left(-\frac{a}{b}\right)^P \nu(c)\nu\left(\frac{d}{h}\right) \quad (6)$$

And

$$\nu\left(\frac{cd}{e}\right)=\left(-\frac{a}{b}\right)^P \nu(d)\nu\left(\frac{c}{e}\right) \quad (7)$$

Substituting the answers for ν (cd), ν (cd/e), and ν (cd/h) into equation (4) yields the following

$$\nu\left(\frac{cd}{eh}\right)=\left(-\frac{a}{b}\right)^P \nu\left(\frac{d}{h}\right)\nu\left(\frac{c}{e}\right) \quad (8)$$

Or

$$\nu(xy)=\left(-\frac{a}{b}\right)^P \nu(x)\nu(y), \quad (9)$$

where x,y are rational and x,y ≥ 1. The answer to (9) can be found by consulting (de Fériet 1977).

$$\nu(x)=\left(-\frac{a}{b}\right)^p x^k \quad (10)$$

Use (**Error! Reference source not found.**) and substitute $\nu(x)=\frac{F(l)}{l}$, *we get*

$$F(l)=\left(-\frac{a}{b}\right)^P l^{1-k}$$

f(L) in I(L) as stated in the assertion, we obtain

$$I(y)=b\left|1-\left[\sum_{\alpha=1}^{n} l_\alpha^{1-k}\right]^{\frac{1}{P}}\right| \quad (11)$$

By applying postulate D from the statement, we may determine the values of the constants b and k. We obtain k=1-P and b=P/(P-1). The proof is now complete.

We refer to the postulate C as Pseudo-additivity.

Theorem 2: *The P-norm information measure* $H_P(l_1,l_2,l_3,\dots,l_n)$ *has the following algebraic properties:*

a) $H_P(l_1,l_2,l_3,\dots,l_n)$ *is symmetric.*
b) H_P *is expansible*
   i.e. $H_P(l_1,l_2,l_3,\dots,l_n,0)=H_P(l_1,l_2,l_3,\dots,l_n)$
c) H_p *is decisive*
   i.e. $H_P(1,0)=H_P(0,1)$
d) H_p *is non additive.*

Proof: From definition 1, properties A through C follow directly, and from Postulate C of Theorem 1, property D follows instantly.

Theorem 3: The algebraic features of the P-norm information measure H_P ($l_1,l_2,l_3,\dots,l_n$) are as follows:

a) $H_P(L) \ge H_P(1,0,\dots..0)=0$
b) $H_P(L) \le H_P(1/n,\dots\dots 1/n)$
c) $H_P(L)$ *is no negative aspects.*
d) $H_P(L)$ *is monotonic function*
e) $H_P(L)$ *is continous in* $\mathbb{R}$
f) $H_P(L)$ *is stable in* $l_\alpha, \alpha = 1,2,\dots\dots,n$
g) $H_P(L)$ *is a concave for every* L *in* Δ_n

Proof: A. We know that

For P > 1, $l_\alpha^P \le l_\alpha \forall \alpha$

$\sum_{\alpha=1}^{n} l_\alpha^P \le \sum_{\alpha=1}^{n} l_\alpha = 1$

$$\left[\sum_{\alpha=1}^{n} l_\alpha^P\right]^{\frac{1}{P}} \le 1 \quad (12)$$

And

$$\frac{P}{P-1} > 0 \quad (13)$$

For 0 < P < 1, $l_\alpha^P \ge l_\alpha \forall \alpha$

$\sum_{\alpha=1}^{n} l_\alpha^P \ge \sum_{\alpha=1}^{n} l_\alpha = 1$

$$\left[\sum_{\alpha=1}^{n} l_\alpha^P\right]^{\frac{1}{P}} \ge 1 \quad (14)$$

And

$$\frac{P}{P-1} < 0 \quad (15)$$

Proof from definition 1 is simple to finish.

B. Lagrange multipliers make it simple to determine that

$$\left[\sum_{\alpha=1}^{n} l_\alpha^P\right]^{\frac{1}{P}} \ge n^{\frac{(1-P)}{P}}, \quad P > 1 \quad (16)$$

And

$$\left[\sum_{\alpha=1}^{n} l_\alpha^P\right]^{\frac{1}{P}} \le n^{\frac{(1-P)}{P}}, \quad 0 < P < 1 \quad (17)$$

The Proof is finished when the values from equations (13) and (15) are used along with the values from equations (16) and (17).

C. Property A provides concrete proof of non negativity.

$D. H_P(L)$ is monotonic iff $H_P(m, 1 - m)$ *is non decreasing on* $m \in \left[0, \frac{1}{2}\right]$

From definition 1, it is clear that

$$H_P(m, 1 - m) = \frac{P}{P - 1}\left[1 - [(1 - m)^P + m^P]^{1/P}\right] \quad (18)$$

Describe a function. G(m) by

$$G(m) = \left[1 - [(1 - m)^P + m^P]^{1/P}\right]$$

Its show that

For P > 1 $\frac{d}{dm}G(m) \geq 0$

And for 0 < P < 1 $\frac{d}{dm}G(m) \leq 0$

So we get

$$\frac{d}{dm}H_P(m, 1 - m) = \frac{P}{P - 1}.\frac{d}{dm}G(m) \quad (19)$$

By using equations (**Error! Reference source not found.**) , (**Error! Reference source not found.**) and (**Error! Reference source not found.**) , we get

$$\frac{d}{dm}H_P(m, 1 - m) \geq 0$$

Hence, $H_P(m, 1 - m)$ is a non-decreasing function and monotonic.

A. $[\sum_{\alpha=1}^{n} l_\alpha^P]^{1/P}$ is a continuous function, as is well known. (see (T. 1953)) for P ∈[0, ∞) So, $H_P(L)$ is continuous function on [0, ∞)

F. We have demonstrated the expansibility of $H_P(L)$ in Theorem 2 B. Consequently, it is evident that

$$H_P(L, 0) = H_P(L) \quad (20)$$

It is clearly that

$$\lim_{t \to 0} H_P(L, e) = H_P(L, 0) = H_P(L)$$

From this and equation (**Error! Reference source not found.**), we say $H_P(L)$ is stable in l_α.

G. It is clear that , From Minkowski inequality (see (Beckenbach and Bellman 2012))

For P > 1,

$$\left[\sum_{\alpha=1}^{n} a_\alpha^P\right]^{1/P} + \left[\sum_{\alpha=1}^{n} b_\alpha^P\right]^{1/P} \geq \left[\sum_{\alpha=1}^{n} (a_\alpha + b_\alpha)^P\right]^{1/P} \quad (21)$$

And for 0 < P < 1,

$$\left[\sum_{\alpha=1}^{n} a_\alpha^P\right]^{1/P} + \left[\sum_{\alpha=1}^{n} b_\alpha^P\right]^{1/P} \leq \left[\sum_{\alpha=1}^{n} (a_\alpha + b_\alpha)^P\right]^{1/P} \quad (22)$$

Put $a_\alpha = \lambda\, l_\alpha$ and $b_\alpha = (1 - \lambda)m_\alpha$, where $0 \leq \lambda \leq 1$, into equation (**Error! Reference source not found.**) and (**Error! Reference source not found.**), we get

For P > 1,

$$\lambda\left[\sum_{\alpha=1}^{n} l_\alpha^P\right]^{1/P} + (1 - \lambda)\left[\sum_{\alpha=1}^{n} m_\alpha^P\right]^{1/P} \geq \left[\sum_{\alpha=1}^{n} (\lambda l_\alpha + (1 - \lambda)m_\alpha)^P\right]^{\frac{1}{P}} \quad (23)$$

And for 0 < P < 1,

$$\lambda\left[\sum_{\alpha=1}^{n} l_\alpha^P\right]^{1/P} + (1 - \lambda)\left[\sum_{\alpha=1}^{n} m_\alpha^P\right]^{1/P} \leq \left[\sum_{\alpha=1}^{n} (\lambda l_\alpha + (1 - \lambda)m_\alpha)^P\right]^{\frac{1}{P}} \quad (24)$$

Using $1 = \lambda + (1 - \lambda)$ and By equations (**Error! Reference source not found.**), (**Error! Reference source not found.**), (**Error! Reference source not found.**), we get

$$\lambda\, H_P(L) + (1 - \lambda)\, H_P(M) \geq H_P(\lambda\, L + (1 - \lambda)\, M) \quad (25)$$

H_P (L) is a concave function in L as a result.

The Proof of Theorem 3 is thus concluded.

3. Final Verdict

The R-Norm Information measure, as described by Boekee and Van der Lubbe (1980), is what I have examined. A few of its attributes are listed and validated. Analytical and algebraic properties have been investigated.

References

[1] Alfred Renyi. 1961. "On Measures of Entropy and Infromation." *Fourth Berkeley Symposium on Mathematical Statistics and Probability* 1: 547–61.

[2] Arimoto, Suguru. 1971. "Information-Theoretical Considerations on Estimation Problems." *Information and Control* 19 (3): 181–94. https://doi.org/10.1016/S0019-9958(71)90065-9.

[3] Beckenbach, Edwin, and Richard Bellman. 2012. "An Introduction To Inequalities." *An Introduction to Inequalities*, 1–2. https://doi.org/10.5948/upo9780883859216.002.

[4] Boekee, D. E., and J. C.A. Van der Lubbe. 1980. "The R-Norm Information Measure." *Information and Control* 45 (2): 136–55. https://doi.org/10.1016/S0019-9958(80)90292-2.

[5] C. E. Shannon. 1948. "A Mathematical Theory of Communication." *The Bell System Technical Journal* 27 (3): 212–14.

[6] Daróczy, Zoltán. 1970. "Generalized Information

Functions." *Information and Control* 16 (1): 36–51. https://doi.org/10.1016/S0019-9958(70)80040-7.

[7] Fériet, J. Kampé de. 1977. "Book Review: On Measures of Information and Their Characterizations." *Bulletin of the American Mathematical Society* 83 (2): 192–97. https://doi.org/10.1090/S0002-9904-1977-14257-2.

[8] Havrda, Jan, and František Charvát. 1967. "Quantification Method of Classification Processes." *Kybernetika* 3 (1): 30–35.

[9] T., E. C. 1953. "Inequalities. By G.H. Hardy, J.E. Littlewood and G. Pólya. 2nd Edition. Pp. Xii, 324. 27s. 6d. 1952. (Cambridge University Press)." *The Mathematical Gazette* 37 (321): 236–236. https://doi.org/10.1017/S0025557200027455.