Vol. 13. Issue.2. 2025 (April-June)

BULLETIN OF MATHEMATICS AND STATISTICS RESEARCH

A Peer Reviewed International Research Journal

http://www.bomsr.com

Email:editorbomsr@gmail.com

INTERNATIONAL STANDARD SERIAL NUMBER 2348-0580

RESEARCH ARTICLE

Correlational Sensitivity among Strictly Bounded Conditional Entropy and Channel Capacity of GIG Function

Rohit Kumar Verma[1], Som Kumari[2]

[1]Associate Professor, Department of Mathematics, Bharti Vishwavidyalaya, Durg, C.G., India

[2]Research Scholar Ph.D. Department of Mathematics, Bharti Vishwavidyalaya, Durg, C.G., India

E-mail- som88sona@gmail.com

DOI:10.33329/bomsr.13.2.44

Article Info

Article Received: 10/05/2025

Article Accepted: 03/06/2025

Published online: 30/06/2025

Abstract

This paper explores the Gaussian Information Gain Function $I(P) = e^{-P^2}$ as an alternative to traditional logarithmic information measures. We define a corresponding entropy function and examine its mathematical properties, including concavity and bounds. Furthermore, we derive a formulation for channel capacity based on Gaussian information, providing a non-logarithmic perspective on information transmission. This approach is particularly relevant in fuzzy systems, uncertain environments, and non-additive information frameworks.

Key words: Gaussian information gain function , Conditional entropy, Weighted entropy, Channel capacity.

1. Introduction

Information theory traditionally measures the amount of information associated with a probabilistic event using the logarithmic function, most notably in Shannon's entropy. For a discrete random variable with probability distribution p_i, the Shannon entropy is defined as:

$$H(p) = -\sum_{i=1}^{n} p_i \log p_i$$

This measure has proven fundamental in the analysis of communication systems, especially in defining the channel capacity the theoretical maximum rate at which information can be reliably transmitted. However, in various complex systems such as those involving fuzzy logic, imprecise probabilities, or non-extensive statistics , alternative measures of information have shown more desirable characteristics. One such alternative is the Gaussian Information Gain Function [4], defined by:

$$I(P) = e^{-P^2}$$

This function measures the information gain from observing an event with probability $p \in [0,1]$. Unlike the logarithmic function, the Gaussian form is bounded, non-logarithmic, and symmetric in nature.It assigns high information content to rare events (low p) and low information content to highly probable events (high p), mimicking the decay of a Gaussian curve.

Now, we define the Gaussian entropy of a discrete distribution p_i as:

$$H_G(p) = \sum_i p_i e^{-p_i^2}$$

This formulation leads to an entropy measure that captures uncertainty differently from Shannon's model. It is especially useful in contexts where the additivity of entropy is not desired, or where small probability differences should result in significant changes in information.

2. New work

2.1 Channel capacity of Gaussian information gain function

In the context of communication systems, the channel capacity is the maximum possible average information gain across a communication channel, constrained by the channel's noise model and input probability distribution. The channel capacity can be formulated as:

$$C_G = \max_{p(x)} \left[\sum_{x,y} p(x,y) e^{-p^2(y/x)} - \sum_y p(y)\, e^{-p^2(y)} \right]$$

This expression reflects the maximum expected difference in Gaussian information gain between the joint distribution *p(x,y)* and the marginal distribution *p(y),* corresponding to mutual information in Shannon's theory.

Let as assume that there is a discrete memory less channel with input and output alphabets of X and Y respectively, input probabilities $P(x)$, output probabilities $P(y)$, and transition probabilities $P(y/x)$, and both X and Y are finite. Reiffen states that the channel is extremely loud if, for all $x \in X$ and $y \in Y$, then we have

$$\frac{P(y)-P(y/x)}{P(y)} = \varepsilon_x(y) \ll 1 \tag{2.1}$$

Equation (2.1) also written as

$$P(y/x) = P(y).(1-\varepsilon_x(y)) \tag{2.2}$$

We observe that

$$\sum_{x\in X} P(x)\,\varepsilon_x(y) = 0 \text{ and } \sum_{y\in Y} P(y)\varepsilon_x(y) = 0 \tag{2.3}$$

with $\varepsilon_x(y) \ll 1$. Later, Susan [3] proposed that $P(y)$ need only be an approximation of the output probabilities rather than the output probabilities itself.

The Gaussian Information Gain Function is

$$I(P) = e^{-P^2} \tag{2.4}$$

The entropy of X is given by

$$H(X) = -\sum_{x\in X} P(x)e^{-P^2(x)} = -\sum_{y\in Y}\sum_{x\in X} P(x,y)e^{-P^2(x)} \tag{2.5}$$

and entropy of Y is given by

$$H(Y) = -\sum_{y\in Y} P(y)e^{-P^2(y)} = -\sum_{x\in X}\sum_{y\in Y} P(x,y)e^{-P^2(y)} \tag{2.6}$$

The Conditional Entropy of X given Y is

$$H(X/Y) = -\sum_{x\in X}\sum_{y\in Y} P(x,y)e^{-P^2(x/y)} \tag{2.7}$$

and Y given X is

$$H(Y/X) = -\sum_{x\in X}\sum_{y\in Y} P(x,y)e^{-P^2(y/x)} \tag{2.8}$$

The joint entropy of X &Y is

$$H(X,Y) = -\sum_{x\in X}\sum_{y\in Y} P(x,y)e^{-P^2(x,y)} \tag{2.9}$$

From equation (2.8),

$$H(Y/X) = -\sum_{x\in X}\sum_{y\in Y} P(x,y)e^{-P^2(y/x)} \tag{2.10}$$

$$H(Y/X) = -\sum_{x\in X}\sum_{y\in Y} P(x)P(y/x)e^{-P^2(y/x)} \tag{2.11}$$

From equation (2.2) we have,

$$P(y/x) = P(y)\big(1-\varepsilon_x(y)\big) \quad where \quad \varepsilon_x(y) \ll 1.$$

Then

$$P^2(y/x) = P^2(y)[1 - 2\varepsilon_x(y) + \varepsilon_x^2(y)]$$

Now,

Apply Taylor expansion to $e^{-P^2(y/x)}$

$$e^{-P^2(y/x)} = e^{-P^2(y)}[1 + 2P^2(y)\varepsilon_x(y) - P^2(y)\varepsilon_x^2(y) + \cdots]$$

Substitute in to Conditional Entropy (2.11)

$$H(Y/X) = -\sum_{x\in X} P(x)\sum_{y\in Y} P(y/x).\,e^{-P^2(y/x)}$$

$$= -\sum_{x\in X} P(x)\sum_{y\in Y} P(y)\big(1-\varepsilon_x(y)\big)\,.\,e^{-P^2(y)}\,[1 + 2P^2(y)\varepsilon_x(y) - P^2(y)\varepsilon_x^2(y) + \cdots]$$

$$= -\sum_{x\in X} P(x)\sum_{y\in Y} P(y)\,.\,e^{-P^2(y)}\,\big[\big(1-\varepsilon_x(y)\big)(1 + 2P^2(y)\varepsilon_x(y) - P^2(y)\varepsilon_x^2(y) + \cdots)\big]$$

$$= -\sum_{x\in X} P(x)\sum_{y\in Y} P(y)\,.\,e^{-P^2(y)}\,[1 + (2P^2(y){-}1)\varepsilon_x(y) - 3P^2(y)\varepsilon_x^2(y) + \cdots)]$$

Total Entropy is

$$H(Y/X) \approx H(Y) + \left[-\sum_{y\in Y} P(y)\,.\,e^{-P^2(y)}\,(2P^2(y){-}1)\sum_{x\in X} P(x)\,\varepsilon_x(y) + \sum_{y\in Y} P(y)\,.\,e^{-P^2(y)}\,3P^2(y)\sum_{x\in X} P(x)\,\varepsilon_x^2(y) - \cdots\right]$$

So, Channel Capacity C would be

$$C = H(Y) - H(Y/X)$$

$$C \approx \left[\sum_{y\in Y} P(y)\,.\,e^{-P^2(y)}\,(2P^2(y){-}1)\sum_{x\in X} P(x)\,\varepsilon_x(y) - \sum_{y\in Y} P(y)\,.\,e^{-P^2(y)}\,3P^2(y)\sum_{x\in X} P(x)\,\varepsilon_x^2(y) + \cdots\right]$$

This is required channel capacity of Gaussian Information Gain function where are true up to the second order of $\varepsilon_x(y) \ll 1$.

2.2 Properties of The Gaussian Information Gain Function.

[P$_1$] The Gaussian information measure $I(P) = e^{-P^2}$ is continuous $\forall\, P \in [0,1]$.

[P$_2$] The higher limit of 1 and the lower limit of e^{-1} define the boundaries of $I\,(P)$.

[P$_3$] The value of $I(P)$ decreases as P increases.

[P$_4$] $H(P) = \sum_{x\in X} Pe^{-P^2(x)}$, the entropy, is a continuous function and a concave function.

Theorem 2.2.1 Let $I(P) = e^{-P^2}$ then for any valid conditional probability distribution [1, 5] $P(y/x)$, the conditional entropy $H(Y/X) = -\sum_{x\in X} P(x)\sum_{y\in Y} P(y/x)\, I(P(y/x))$ is bounded i,e;

$$H(Y/X) \in \left[-\sum_{x\in X} P(x)\sum_{y\in Y} P(y/x)\,.\,e^{-1}, -\sum_{x\in X} P(x)\sum_{y\in Y} P(y/x)\,.\,1\right]$$

Proof: Since $P(y/x) \in [0,1], then\ .P^2 \in [0,1]$.

Thus, $I(P) = e^{-P^2} \in [e^{-1}, 1]$

Therefore, for any value of $P(y/x)$

$$e^{-1} \le I(P(y/x)) \le 1$$

$$-\sum_{x\in X} P(x)\sum_{y\in Y} P(y/x)\, e^{-1} \le -\sum_{x\in X} P(x)\sum_{y\in Y} P(y/x)\, I(P(y/x)) \le -\sum_{x\in X} P(x)\sum_{y\in Y} P(y/x)\,.\,1$$

i.e; $-\sum_{x\in X} P(x)\sum_{y\in Y} P(y/x)\, e^{-1} \le H(Y/X) \le -\sum_{x\in X} P(x)\sum_{y\in Y} P(y/x)\,.\,1$

Thus $H(Y/X) \in \left[-\sum_{x\in X} P(x)\sum_{y\in Y} P(y/x)\,.\,e^{-1}, -\sum_{x\in X} P(x)\sum_{y\in Y} P(y/x)\,.\,1\right]$

Hence $H(Y/X)$ is bounded.

Theorem 2.2.2 Let $I(P) = e^{-P^2}$ then the entropy [2] defined by $H(X) = -\sum_{x\in X} P(x)\, I\big(P(x)\big)$ is non – additive, i.e; $H(X,Y) \neq H(X) + H(Y)$

Proof: In Classical (Shannon) entropy, additivity holds when X & Y are independent.

$$H(X,Y) \neq H(X) + H(Y)$$

$$H(X,Y) = -\sum_{x,y} P(x)P(y)I\big(P(x)P(y)\big)$$

$$= -\sum_{x,y} P(x)P(y)e^{-P^2(x)P^2(y)}$$

While $\quad H(X) + H(Y) = -\sum_{x\in X} P(x)e^{-P^2(x)} - \sum_{x\in X} P(y)e^{-P^2(y)}$

Since $\quad e^{-P^2(x)P^2(y)} \neq e^{-P^2(x)}.e^{-P^2(y)}$

It follows that $\quad H(X,Y) \neq H(X) + H(Y)$

Thus, Entropy is non-additive.

Theorem 2.2.3 Let $P(y/x) = P(y)(1-\varepsilon_x(y))$ where $\varepsilon_x(y) \ll 1$.Then the entropy

$$H(Y/X) = -\sum_{x\in X}\sum_{y\in Y} P(x)P(y/x)e^{-P^2(y/x)}$$

is sensitive to such correlations and can be approximated via. Taylor expansion.

Proof: Assume $P(y/x) = P(y)(1-\varepsilon_x(y))$ where $\varepsilon_x(y) \ll 1$.

Then

$$P^2(y/x) = P^2(y)[1 - 2\varepsilon_x(y) + \varepsilon_x^2(y)]$$

Using Taylor expansion

$$e^{-P^2(y/x)} = e^{-P^2(y)}[1 + 2P^2(y)\varepsilon_x(y) - P^2(y)\varepsilon_x^2(y) + \cdots \ldots\ldots]$$

Then

$$H(Y/X) \approx -\sum_{x\in X} P(x)\sum_{y\in Y} P(y/x)\,.\,e^{-P^2(y)}\,[1 + 2P^2(y)\varepsilon_x(y) - P^2(y)\varepsilon_x^2(y) + \cdots \ldots\ldots]$$

Thus small $\varepsilon_x(y)$ effects entropy up to second order showing correlation sensitivity.

3. Conclusion

In this paper, we have explored the entropy properties of the Gaussian Information Gain (GIG) function defined as $I(P) = e^{-P^2}$ .Unlike conventional information measures such as Shannon or Renyi entropy, the GIG function introduces a smooth, rapidly decaying behavior for probability weights, leading to novel and bounded entropy characteristics. We established three key theorems highlighting distinct aspects of this measure:

1. Bounded Conditional Entropy: We proved that the entropy derived from the GIG function is strictly bounded, which ensures numerical stability and predictability in uncertainty estimation.
2. Non-Additivity: Unlike additive measures such as Shannon entropy, the GIG entropy exhibits non-additivity, making it sensitive to the joint distribution structure and ideal for modeling correlated systems.
3. Correlation Sensitivity: We demonstrated through second-order approximation that the GIG-based entropy is highly responsive to small perturbations in conditional distributions, suagestina its utility in contexts where fine-grained differences in correlation need to be captured.

Additionally, we derived the channel capacity based on this function using a novel entropy maximization approach under Gaussian constraints.This derivation not only extends traditional information theory but also opens up new avenues in communication system design, particularly in uncertain or non-traditional environments. Overall, the Gaussian Information Gain function presents a rich and flexible framework for information representation, offering theoretical and practical advantages for modern inference, communication, and decision-making systems.

References

[1]. Gallager, R. (1965). A simple derivation of the coding theorem and some applications. *IEEE Transactions on Information Theory*, *11*(1), 3–18.

[2]. Rényi, A. (1961). On measures of entropy and information. In *Proceedings of the Fourth Berkeley Symposium on Mathematical Statistics and Probability* (Vol. 1, pp. 547–561). University of California Press.

[3]. Shannon, C. E. (1948). A mathematical theory of communication. *Bell System Technical Journal*, *27*(3), 379–423; *27*(4), 623–659.

[4]. Susan, S., & Hanmandlu, M. (2013). A non-extensive entropy feature and its application to texture classification. *Neurocomputing*, *120*, 214–225.

[5]. Tsallis, C., Mendes, R. S., & Plastino, A. R. (1998). The role of constraints within generalized non-extensive statistics. *Physica A: Statistical Mechanics and its Applications*, *261*(3–4), 534–554.