Mathematics and Statistics 12(3): 283-291, 2024 http://www.hrpub.org
DOI: 10.13189/ms.2024.120308

On Use of Entropy Function for Validating Differential Calculus Results

Omdutt Sharma[1,*], Surender Kumar[2], Naveen Kumar[2], Pratiksha Tiwari[3]

[1]Department of Mathematics, P.D.M. University, India
[2]Department of Mathematics, Baba Mast Nath University, India
[3]Delhi Institute of Advanced Studies, India

Received March 5, 2024; Revised April 26, 2024; Accepted May 26, 2024

Cite This Paper in the Following Citation Styles

(a): *[1] Omdutt Sharma, Surender Kumar, Naveen Kumar, Pratiksha Tiwari , "On Use of Entropy Function for Validating Differential Calculus Results," Mathematics and Statistics, Vol. 12, No. 3, pp. 283 - 291, 2024. DOI: 10.13189/ms.2024.120308.*

(b): *Omdutt Sharma, Surender Kumar, Naveen Kumar, Pratiksha Tiwari (2024). On Use of Entropy Function for Validating Differential Calculus Results. Mathematics and Statistics, 12(3), 283 - 291. DOI: 10.13189/ms.2024.120308.*

Abstract Rolle's Theorem (RT) and Lagrange's Mean-value Theorem (LMVT) are significant for pure and applied mathematics, and they have applications in various other fields such as management, physics etc. RT is significant in finding the projectile trajectory's maximum height and in information theory, and the entropy function (measure) is used to measure the uncertainty of information. RT is used to analyze the graphs of annual performance in any field. Since information is necessary to analyze any performance and in information theory, entropy measure is a significant tool to quantize the uncertainty so by using the concept of RT and LMVT in information theory the uncertainty and vagueness or noise can be minimized or maximized. In this manuscript, the concept of differential calculus, i.e., RT and LMVT are used for validation of the entropy function. In this paper, characteristics of differential calculus in information entropy function have been discussed. It has been shown that the entropy function satisfies RT and LMVT. It also describes the conditions when Rolle's Theorem becomes the necessary and sufficient condition for entropy function. Theorems are proved related to the concept of differential calculus in information theory which shows that by using the existing entropy function some new entropies can be derived.

Keywords Entropy Function, Rolle's Theorem, Lagrange Mean-value Theorem, Cauchy Mean-value Theorem, Probabilistic Entropy

1. Introduction

In recent years as a promising movement, the principles of information theory have been exploited to unravel the dilemmas related to the information. In the late 1940s after the initiation by Shannon and Wiener, information theory was operated to analyze the basic measure of reliably transmitting messages via a noisy communication channel. So far, information theory has created a deep effect on various areas involving electrical engineering, computer science, mathematics, physics, and social sciences.

In the 19th century, the notion of invented entropy was related to thermodynamics where it was closely associated with heat flow and central to the 2nd law of thermodynamics. After that Gibbs [1] to measure the disorder of the large dynamical system under molecular collections in thermodynamics, entropy measures were applied. Fisher [2] utilized this notion in the expansion of the fundamentals of theoretical statistics. Afterward, in the late 1940s, Claude Shannon [3], an electrical engineer and mathematician indicated the association between entropy and information content. In the 1950s, this notion was further utilized by Kullback and Leibler [4]. After that to develop the generalization of Shannon's entropy a lot of work was carried out by researchers like Renyi [5], Kapur [6], Tsallis [7], and others. Sharma et al. [8] suggested entropy for fuzzy soft matrices and apply in data reduction applications, Sharma and Gupta [9], by using conic sections derived the entropy measures also discussed their application.

Since our daily life's dilemmas are full of inexact and vague information. Thus, to solve these types of problems probabilistic and statistical tools are not the finest existing gizmo. Thus, to handle the imprecise and vague information, Zadeh [10] introduced a non-probabilistic approach known as the fuzzy set. The entropy measure of a fuzzy set is a measure of fuzziness; such a measure illustrates the sharpness of the membership function of the element of a fuzzy set which is not essentially related to the random experiments. Zadeh [11] defined the entropy measure for a fuzzy set as the weighted Shannon's entropy. Such measures can be inferred as the uncertainty related to the fuzzy event. The first non-probabilistic entropy measure was initiated by De Luca and Termini [12], who defined the entropy using Shannon's entropy. Later Bhandari and Pal [13], Sharma et al. [14], and other researchers introduced different non-probabilistic entropy measures.

In differential calculus, according to RT at a certain point between two different points having the same value where the 1st order differentiation of a real-valued differentiable function is zero,Bhaskara (II) (1114-1185) recognized the Rolle's Theorem Selin [15]. But it is named after 'Michel Rolle' because 'Rolle's' [16],[17] prove it considering the case of polynomial functions (Aczel and Doroczy [18], Mathai and Rathie [19]. Since proof did not utilize the methods of differential calculus, hence it was considered fallacious at that time. In 1823 Cauchy proved the theorem first time as a corollary of a proof of the MVT. "Moritz Wilhelm Drobisch" of Germany in 1834 and "Giusto Bellavitis" of Italy in 1846 were the first one used the name "Rolle's Theorem" (Cajori,[20]). Also some other researchers work in the field of entropy or uncertainty measures and their applications [21], [22], [23], [24], [25].

The rest of paper is arranged as: Section 2 deals with the related concept in which related concepts like probabilistic entropy, non-probabilistic entropy; RT, and the Geometrical interpretation of RT are discussed in brief. Section 3 deals with the proposed work in which it proved that the entropy function satisfies Rolle's Theorem and also states that it satisfies also Lagrange's mean value theorem and under what condition satisfies Cauchy's Mean value theorem. The geometrical interpretation and physical significance of LMVT in information entropy are also discussed in the same section. Section 4 deals with discussion. At last, the paper is concluded.

2. Preliminaries

Shannon defined entropy for probability distribution $p_x \in P$ as shown:

For discrete case:

$$H(P) = -\sum_{p_x \in P} p_x \log p_x$$

where $0 \leq p_x \leq 1 \ and \ \sum_{p_x \in P} p_x = 1$

For continuous case:

$$H(P) = \int_0^1 p_x \log p_x \ dp_x$$

The assumed axioms for a valid entropy measure are as follows:

I. When the probabilities are uniformly distributed, the entropy should be a monotonically increasing function of a number of outcomes.
II. When the category of outcomes splits into subcategories by anyone, then the new entropy of the extended system equals to the sum of the old system entropy, also the new entropy weighted by its probability creation of more choice increases uncertainty weighted according to the likelihood of the increased choice being relevant.
III. Uncertainty should be a continuous function of the p_x and the entropy function satisfies the following properties:

- The entropy measure is non-negative.
- The entropy function is symmetric in the nature of its arguments.
- The entropy function attains maximum value when$p_x = 0.5 \ or \frac{1}{2} \ or \ all \ p_x \ are \ equal.$
- Entropy function is a continuous function for all p_x.
- Entropy value for certain outcomes is minimum or possibly zero.
- Entropy remains same after the addition of probability of impossible outcome.

Corresponding to Shannon's entropy "DE Luca and Termini" [1972] introduced fuzzy entropy which is non-probabilistic in nature as:

$$H(Z) = -\sum_{i=1}^{n} [\mu_Z(x_i) \log \mu_Z(x_i) + (1 - \mu_Z(x_i)) \log(1 - \mu_Z(x_i))]$$

where Z is a fuzzy set and $x_i \in Z$ and $\mu_Z(x_i)$ is the membership value of x_i in Z and lies in the interval $[0, 1]$.

The conditions shown below are satisfied by valid fuzzy entropy:

- Fuzzy entropy is greater than and equal to zero.
- Fuzzy entropy is the minimum or possible zero corresponding to a crisp set.
- Fuzzy entropy is the maximum for most fuzzy sets.
- $H(Z^*) \leq H(Z)$, where Z^* is a sharpened version of A.
- $H(Z) = H(Z^c)$, where Z^c is the complement of Z.

Rolle's Theorem: RT is a special case of the MVT which is significant in the proof of the fundamental theorem of differential calculus. The theorem states as follows:

"Let f be a real-valued function defined in the closed interval $[a,b]$ such that

- f is continuous on $[a,b]$;
- f is differentiable on (a,b);
- $f(a)=f(b)$.

Then, there exists a real number c in the open interval (a,b) such that $f'(c)=0$."

Remark: When any function satisfies all the three conditions stated above then RT is applicable to that function.

Geometrical significance of RT:

Let Rolle's Theorem be applicable on f defined on $[a,b]$. Then, continuity of f on $[a,b]$, follows that a graph of $f(x)$ can be drawn in range of $x=a$ and $x=b$.

Also, differentiability of $f(x)$ in open interval of 'a' and 'b' means that $f(x)$ has a tangent at each point of (a,b).

Now, the existence of $c\in(a,b)$ having $f'(c)=0$ indicates that the tangent at $x=c$ has a slope 0, which means that the tangent at c is parallel to the x-axis.

3. Proposed Work

In this section, the relation between entropy function and differential calculus is discussed in the form of theorems:

Theorem 1: Every probabilistic entropy function (measure) satisfies RT.

Or

Theorem: RT is a necessary condition for a valid probabilistic entropy function (measure).

Proof: Let $H(p)$ be a valid probabilistic entropy function defined on a probability distribution P, to prove theorem, it shows that entropy function satisfies all Rolle's Theorem conditions defined as:

- "f is continuous on $[a,b]$;
- f is differentiable on (a,b);
- $f(a)=f(b)$.
- Then, $\exists, c\in(a,b)$ such that $f'(c)=0$."

Since $H(p)$ is valid entropy so it is continuous on $[0,1]$, and also $H(p)=0$ or minimum at $p(x)=0\ and\ 1$. $H(p)$ is also a symmetric function so $H(0)=H(1)$. Every probabilistic entropy function is differentiable in $(0,1)$. Also, in a probabilistic entropy function $\exists, c\in(0,1)\ and\ c=0.5$ such that $f'(c)=0$, which is the point of most uncertainty. Since all conditions of Rolle's Theorem are satisfied hence theorem is proved.

Corollary 1: Every probabilistic entropy function (measure) satisfies Lagrange's Mean-value Theorem.

Proof: Let $H(p)$ be a valid probabilistic entropy function defined on a probability distribution P, to prove the theorem it shows that the entropy function satisfies all LMV theorem conditions defined as:

- "f is continuous on $[a,b]$;
- f is differentiable on (a,b);
- $f(a)=f(b)$.
- Then, there exists a point c in the open interval (a,b) such that $f'(c)=\frac{f(b)-f(a)}{b-a}$."

Since $H(p)$ is valid entropy so it is continuous on $[0,1]$, and also $H(p)=0$ or minimum at $p(x)=0\ and\ 1$. $H(p)$ is also a symmetric function so $H(0)=H(1)$. Every probabilistic entropy function is differentiable in $(0,1)$.

Let $H(p_1)$ and $H(p_2)$ be the entropy with respect to probabilities $p_1\ \&\ p_2$ and these points form the equation of the secant line joining $(p_1,H(p_1))$ and $(p_2,H(p_2))$, which is

$$h(p)=\frac{H(p_2)-H(p_1)}{p_2-p_1}(p-p_1)+H(p_1)$$

The vertical difference between the graph of H(p) and h(p) is

$$E(p)=H(p)-h(p)$$

$$=H(p)-H(p_1)-\frac{H(p_2)-H(p_1)}{p_2-p_1}(p-p_1)$$

Since $H(p)$ is continuous on [0, 1] and differentiable on (0, 1) so is $E(p)$. Moreover, $E(p_1)=0$ and $E(p_2)=0$; so $E(p)$ satisfies the hypothesis of RT on [0, 1]. Thus, there is a point q in (0, 1) such that $E'(q)=0$ but we have

$$E'(p)=H'(p)-\frac{H(p_2)-H(p_1)}{p_2-p_1}$$

So,

$$E'(q)=H'(q)-\frac{H(p_2)-H(p_1)}{p_2-p_1}$$

Since $E'(q)=0$, we have

$$H'(q)=\frac{H(p_2)-H(p_1)}{p_2-p_1}$$

Since all conditions of LMVT are satisfied hence corollary is proved.

Theorem 2: Every fuzzy entropy function (measure) satisfies RT.

or

Theorem: RT is a necessary condition for a valid fuzzy entropy function (measure).

Proof: Let $H(A)$ be a valid fuzzy entropy function defined on a fuzzy set A, to prove the theorem it shows that $H(A)$ satisfies all Rolle's Theorem conditions defined as:

- f is continuous on $[a,b]$;
- f is differentiable on (a,b);
- $f(a)=f(b)$.
- Then, there exists a real number c in the open interval (a,b) such that $f'(c)=0$.

Since $H(A)$ is valid fuzzy entropy so it is continuous on $[0,1]$, and also $H(A)=0$ or minimum at $\mu_A(x_i)=0\ and\ 1$. $H(A)$ is also a symmetric function so $H(0)=H(1)$. Every fuzzy entropy function is differentiable in

$(0, 1)$. Also, in fuzzy entropy function there exists a real number $c = 0.5$ in open interval $(0, 1)$ such that $f'(c) = 0$, which is the point of most fuzziness. Since all conditions of Rolle's Theorem are satisfied hence theorem is proved.

Corollary 2: Every fuzzy entropy function (measure) satisfies Lagrange's Mean-value Theorem.

Proof: Let $H(A)$ be a valid fuzzy entropy function defined on membership function $\mu_A(x)$, to prove the theorem it shows that the entropy function satisfies all LMV theorem conditions stated above:

Since $H(A)$ is valid entropy so it is continuous on $[0, 1]$, and also $H(A)$ minimum at $\mu_A(x) = 0\ and\ 1$. $H(A)$ is also a symmetric function so $H(0) = H(1)$. Every fuzzy entropy function is differentiable in$(0, 1)$.

Let $H(A_1)$ and $H(A_2)$ be the entropy with respect to probabilities $\mu_A(x_1)$ & $\mu_A(x_2)$ and these points form the equation of the secant line joining $(\mu_A(x_1),\ H(A_1))$ and $(\mu_A(x_2), H(A_2))$, which is

$$h(A) = \frac{H(A_2) - H(A_1)}{\mu_A(x_2) - \mu_A(x_1)}(\mu_A(x) - \mu_A(x_1)) + H(A_1)$$

the vertical difference between the graph of H(A) and h(A) is

$$E(A) = H(A) - h(A)$$

$$= H(A) - H(A_1) - \frac{H(A_2) - H(A_1)}{\mu_A(x_2) - \mu_A(x_1)}(\mu_A(x) - \mu_A(x_1))$$

Since $H(A)$ is continuous on [0, 1] and differentiable on (0, 1) so is $E(A)$. Moreover, $E(\mu_A(x_1)) = 0$ and $E(\mu_A(x_2)) = 0$; so $E(A)$ satisfies R.T on [0, 1]. Thus, there is a point $\mu_A(x_3)$ in (0, 1) such that $E'(\mu_A(x_3)) = 0$ but we have,

$$E'(A) = H'(A) - \frac{H(\mu_A(x_2)) - H(\mu_A(x_1))}{\mu_A(x_2) - \mu_A(x_1}$$

So,

$$E'(\mu_A(x_3)) = H'(\mu_A(x_3)) - \frac{H(\mu_A(x_2)) - H(\mu_A(x_1))}{\mu_A(x_2) - \mu_A(x_1}$$

Since $E'(\mu_A(x_3)) = 0$, we have6

$$H'(\mu_A(x_3)) = \frac{H(\mu_A(x_2)) - H(\mu_A(x_1))}{\mu_A(x_2) - \mu_A(x_1}$$

Since all conditions of LMVT are satisfied hence corollary is proved.

Remark: It is not necessary that every function that satisfies Rolle's theorem and LMVT is a entropy function (measure).

Theorem 3: Every probabilistic and non-probabilistic entropy function is neither increasing nor decreasing on $(0, 1)$.

Proof: Since all entropy function is strictly increasing and decreasing on $(0, 0.5)$ and $(0.5, 1)$ respectively. Thus, all entropy function is neither increasing nor decreasing on$(0, 1)$.

Theorem 4: Every entropy function (measure) with respect to another entropy function does not satisfy Cauchy's Mean-value Theorem.

Proof: Let $H(x)$& $E(x)$ be two entropy functions defined on $[0, 1]$. Since both functions are continuous on $[0, 1]$ and differentiable on $(0, 1)$. But for both functions $H(0) = H(1)$ and $E(0) = E(1)$, hence the condition for the denominator function is not satisfied. Hence entropy function with respect to another function does not satisfy Cauchy's Mean-value theorem.

Corollary 3: Every entropy function with respect to another function $g(x)$ satisfies Cauchy's Mean-value Theorem if $g(x)$ is continuous on $[0, 1]$ and differentiable on $(0, 1)$ and if $g'(x) \neq 0$ for any $x \in (0,1)$. Then there exists at least one $c \in (0, 1)$ such that

$$\frac{H(1) - H(0)}{g(1) - g(0)} = \frac{H'(c)}{g'(c)}$$

Proof: Define function

$$E(x) = H(x) - \left(\frac{H(1) - H(0)}{g(1) - g(0)}\right)g(x)$$

Since $H'(x)$ and $g(x)$ are continuous in [0,1].
Therefore $E(x)$ is also continuous on [0,1].
Again since $H'(x)$ and $g'(x)$ exist in (0,1).
Therefore $E'(x)$ also exists in (0,1) and is equal to

$$H'(x) - \left(\frac{H(1) - H(0)}{g(1) - g(0)}\right)g'(x)$$

Clearly $E(0) = E(1). Thus\ E(x)$ satisfies all the conditions of RT. Then ∃, at least one value $c \in (0,1)$, such that $E'(c) = 0$.

That is $0 = H'(c) - \left(\frac{H(1)-H(0)}{g(1)-g(0)}\right)g'(c)$ or $H'(c) = \left(\frac{H(1)-H(0)}{g(1)-g(0)}\right)g'(c)$

Theorem 5: A non-negative real-valued function defined in [0, 1] is an entropy function if and only if it is strictly increasing and decreasing on $[0, 0.5]$ and $[0.5, 1]$ respectively and also satisfies RT.

OR

Theorem: A non-negative real-valued function that is defined in [0,1] and satisfies Rolle's Theorem is an entropy function if and only if it is strictly increasing and decreasing on $[0, 0.5]$ and $[0.5, 1]$ respectively.

Geometrical Interpretation: Here geometrically the RT and LMVT are interpreted for entropy function. As shown in Fig. 1: the entropy function curve is continuous in [0, 1] and differentiable in (0, 1) and $H(0) = H(1)$. There also exists a point $c = 0.5$, such that $f'(c) = 0$ and $f(c) = maximum$ where $'c'$ lies in interval $(0, 1)$, from Figure 1. It is shown that the graph of $H(x)$ has a tangent at each point of (0, 1).

Also, we state that all entropy function satisfies LMVT, because LMVT states as:

"Let f be a real function such that f is continuous on

$[0,1]$ and differentiable on$(0,1)$. Then, $\exists$ a real number $c \in (0,1)$ s.t $f'(c) = \left(\frac{f(1)-f(0)}{(1-0)}\right)$".

Now, if we take the points $A[0.2, H(0.2)]$ and $B[0.8, H(0.8)]$ on above curve then slope of chord $AB = \frac{H(0.8)-H(0.2)}{0.8-0.2}$, thus from above discussion, we have $H'(c) = slope\ of\ chord\ A$.

This shows that the tangent at $x = c$ is parallel to AB and if $H(a) = H(b)$ then they are also parallel to X- axis.

"Physical Significance of the Mean value theorem in Information Entropy": Let $H(a) and\ H(b)$ be the uncertainty values at the $a\ \&\ b$ two different probability distributions (or membership value) respectively.

Then average uncertainty of information = $\frac{H(b)-H(a)}{(b-a)}$.

Also, $H'(c)$ denotes the instantaneous uncertainty of information at the probability (membership value) c.

By the Mean-value theorem: $H'(c) = \left(\frac{H(b)-H(a)}{(b-a)}\right)$.

So the mean-value theorem says that at some probability (or membership value) c between $a\ \&\ b$, then instantaneous uncertainty of information would be equal to the average uncertainty.

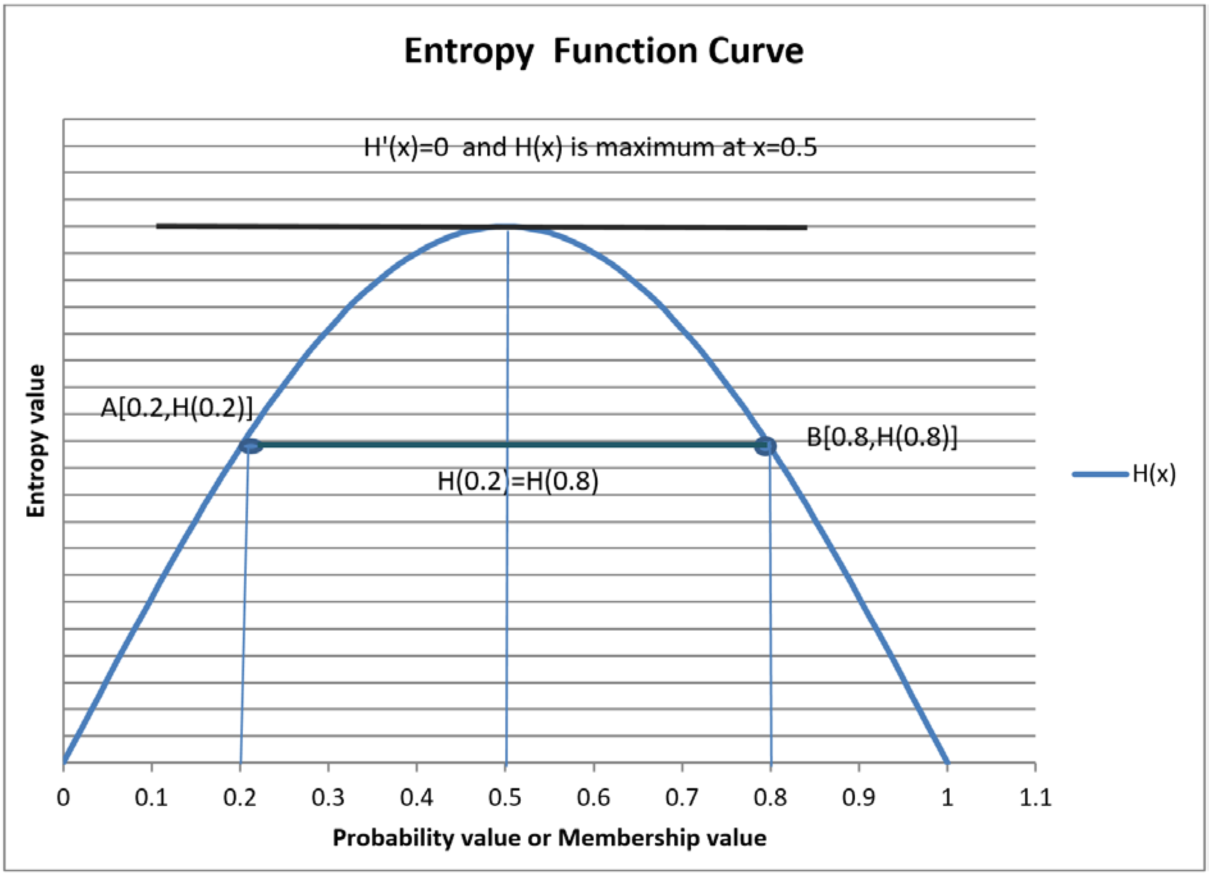

Figure 1. Plot of General Entropy Function

Some Properties of Entropy Function:

Here some properties of entropy function are discussed in the form of theorem.

Theorem 6: Addition of two probabilistic Entropy functions is entropy function.

Proof: Let $F(p)$ & $G(p)$ be two Probabilistic entropy functions and $H(p)$ be another function defined in the same domain as

$$H(p) = F(p) + G(p).$$

Now $H(p)$ is probabilistic entropy function if it satisfies the condition which is stated above in section 2.

1. $H(p)$ is non-negative, since the addition of two non-negative functions is non-negative.
2. $H(p)$ is symmetric in nature with respect to its argument because addition of two symmetric functions is symmetric with respect to its arguments.
3. $H(p)$ attains maximum value when $p_x = 0.5\ or\ \frac{1}{2}\ or\ all\ p_x\ are\ equal.$
4. $H(p) = 0$ when $p = 0\ or\ 1.$
5. $H(p)$ is a continuous function for all p_x because addition of two continuous functions is continuous.
6. The value of $H(P)$ should not change if an impossible outcome is added to the probability scheme.

Since $H(P)$ satisfies all the conditions of valid probabilistic entropy measures so it is a valid probabilistic entropy measure.

Corollary 4: Addition of fuzzy entropy function is fuzzy entropy.

Proof: Let $F(t)$ & $G(t)$ be two fuzzy entropy functions and $H(t)$ be another function defined in the same domain as

$$H(t) = F(t) + G(t).$$

Now $H(t)$ is fuzzy entropy function if it satisfies the condition of valid fuzzy entropy i.e;

1. $H(t)\ is\ equal\ to\ zero$, when t is crisp since $F(t) = G(t)$ is equal to zero when is crisp.
2. Now $H(t)$ is equal to $H(\bar{t})$;
   $H(t)$ is equal to$F(t) + G(t)$
   This implies $H(\bar{t})is\ equal\ to$
   $F(\bar{t}) + G(\bar{t})$, since$F(t)$ and $G(t)$ are valid entropy so $F(\bar{t})$ is equal to $F(t)$ and $G(\bar{t})$ is equal to $G(t)$, thus $H(\bar{t}) = F(t) + G(t) = H(t)$
3. $H(t)$ attains maximum value if $\mu_t(x) = 0.5\ \forall x \in t.$ Since $F(t)$ & $G(t)$ attain maximum value at $\mu_t(x) = 0.5\ \forall x \in t$ due to valid fuzzy entropy hence their addition also attains maximum value.
4. Since $H(t)$ is equal to $F(t) + G(t)$
   Then $H(t^*)$ is equal to $F(t^*) + G(t^*)$, As F & G are valid fuzzy entropy so $F(t)\ is\ greater\ than\ equal\ to\ \ F(t^*)$ and $G(t) \geq G(t^*)$ therefore $H(t) \geq F(t^*) + G(t^*) = H(t^*)$
   $$H(t) \geq H(t^*)$$

Since $H(t)$ satisfies all the conditions of valid fuzzy entropy measures so it is a valid entropy.

Theorem 7: Multiplication of two probabilistic entropy functions is entropy function.

Proof: Let $F(p)$ & $G(p)$ be two Probabilistic entropy functions and $H(p)$ be another function defined in the same domain as

$$H(p) = F(p) * G(p).$$

Now $H(p)$ is probabilistic entropy function if it satisfies the condition which is stated above in section 2.

1. $H(p)$ is non-negative, since the multiplication of two non-negative functions is non-negative.
2. $H(p)$ is symmetric in nature with respect to its argument because multiplication of two symmetric functions is symmetric with respect to its arguments.
3. $H(p)$ attains maximum value when $p_x = 0.5\ or\ \frac{1}{2}\ or\ all\ p_x\ are\ equal.$
4. $H(p) = 0$ when $p = 0\ or\ 1.$
5. $H(p)$ is a continuous function for all p_x because multiplication of two continuous functions is continuous.
6. The value of $H(P)$ remains same after the addition of probability of impossible outcome.

Since $H(P)$ satisfies all the conditions of valid probabilistic entropy measures so it is a valid probabilistic entropy measure.

Corollary 5: Multiplication of fuzzy entropy function is fuzzy entropy.

Proof: Let $F(u)$ & $G(u)$ be two fuzzy entropy functions and $H(u)$ be another function defined in the same domain as

$$H(u) = F(u) * G(u).$$

Now $H(u)$ is a fuzzy entropy function if it satisfies the condition of valid fuzzy entropy i.e;

1. $H(u) is\ equal\ to\ zero$, when u is crisp since $F(u) = G(u)$ is equal to zero, when A is crisp.
2. Now $H(u)\ is\ equal\ to\ H(\bar{u})$;
   $H(u)$ is equal to $F(u) * G(u)$
   $=> H(\bar{u}) = F(\bar{u}) * G(\bar{u})$, since $F(u)$ & $G(u)$ are valid entropy so $F(\bar{u})\ is\ equal\ to\ F(u)\ and\ G(\bar{u})$is equal to $G(u)$, thus $H(\bar{u}) = F(u) * G(u) = H(u)$
3. $H(u)$ attains maximum value if $\mu_u(x) = 0.5\ \forall x \in u.$ Since $F(u)$ & $G(u)$ attain maximum value at $\mu_u(x) = 0.5\ \forall x \in u$ due to valid fuzzy entropy hence their multiplication also attains maximum value.
4. Since $H(u)$ is equal to $F(u) * G(u)$
   Then $H(u^*)$ is equal to $F(u^*) * G(u^*)$, As F & G are valid fuzzy entropy so $F(u)$ is greater than and equal to $F(u^*)$ and $G(u)$ is greater than and equal to $G(u^*)$ therefore $H(u) \geq F(u^*) * G(u^*) = H(u^*)$
   $$H(u) \geq H(u^*)$$

Since $H(u)$ satisfies all the conditions of valid fuzzy entropy measures so it is a valid entropy.

Theorem 8: Subtraction of two probabilistic entropy functions $F(p)$ & $G(p)$ is an entropy function if $F(p) > G(p)$

Proof: Let $H(p)$ be a function defined as $H(p) = F(p) - G(p)$ where $F(p)$ & $G(p)$ be two Probabilistic entropy functions s.t $F(p) > G(p)$.

Now $H(p)$ is a probabilistic entropy function if it satisfies the condition which is stated above in section 2.

1. $H(p)$ is non-negative, by definition.
2. $H(p)$ is symmetric in nature with respect to its argument because subtraction of two symmetric function is symmetric with respect to its arguments.
3. $H(p)$ attains maximum value when $p_x = 0.5\ or \frac{1}{2}\ or\ all\ p_x\ are\ equal.$
4. $H(p) = 0$ when $p = 0\ or\ 1$.
5. $H(p)$ is a continuous function for all p_x because subtraction of two continuous functions is continuous.
6. The value of $H(P)$ remains same after the addition of probability of impossible outcome.

Since $H(P)$ satisfies all the conditions of valid probabilistic entropy measures so it is a valid probabilistic entropy measure.

Remark: The above result is true only when a smaller valued entropy function is subtracted from a greater valued entropy function.

Corollary 6: Subtraction of fuzzy entropy function $F(v)$ & $G(v)$ is an entropy function if $F(v) > G(v)$ is fuzzy entropy.

Proof: If $F(v)$ & $G(v)$ be two fuzzy entropy functions and $H(v)$ be another function defined in the same domain as

$$H(v) = F(v) - G(v).$$

Now $H(v)$ is a fuzzy entropy function if it satisfies the condition of valid fuzzy entropy i.e;

1. $H(v)\ is\ equal\ to\ zero$, when v is crisp, since $F(v) = G(v)$ is equal to zero, when v is crisp.
2. Now $H(v)$ is equal to $H(\bar{v})$;
   $H(v)\ is\ equal\ to\ F(v) - G(v)$
   $=> H(\bar{v})\ is\ equal\ to\ F(\bar{v}) - G(\bar{v})$, since $F(v)$ & $G(v)$ are valid entropy so $F(\bar{v})$ is equal to $F(v)$ and $G(\bar{v})$ is equal to $G(v)$, thus $H(\bar{v}) = F(v) - G(v) = H(v)$
3. $H(v)$ attains maximum value if $\mu_v(x) = 0.5\ \forall x \in v$. Since $F(v)$ & $G(v)$ attain maximum value at $\mu_v(x) = 0.5\ \forall x \in v$ due to valid fuzzy entropy hence their subtraction also attains maximum value.
4. Since $H(v)$ is equal to $F(v) - G(v)$, then $H(v^*)$ is equal to
   $F(v^*) - G(v^*)$, As F and G are valid fuzzy entropy so $F(v) \geq F(v^*)$ and $G(v)$ are greater than and equal to $G(v^*)$ therefore $H(v) \geq F(v^*) - G(v^*) = H(v^*)$. Hence $H(v)$ is greater than and equal to $H(v^*)$

Since $H(v)$ satisfies all the conditions of valid fuzzy entropy measures so it is a valid entropy.

Theorem 9: Division of two probabilistic entropy functions is an entropy function.

Proof: Let $F(p)$ & $G(p)$ be two Probabilistic entropy functions and $H(p)$ be another function defined in the same domain as

$$H(p) = F(p)/G(p).$$

Now $H(p)$ is a probabilistic entropy function if it satisfies the condition which is stated above in section 2.

1. $H(p)$ is non-negative, since the division of two non-negative functions is non-negative.
2. $H(p)$ is symmetric in nature with respect to its argument because the division of two symmetric functions is symmetric with respect to its arguments.
3. $H(p)$ attains maximum value when $p_x = 0.5\ or \frac{1}{2}\ or\ all\ p_x\ are\ equal.$
4. $H(p) = 0$ when $p = 0\ or\ 1$.
5. $H(p)$ is a continuous function for all p_x because division of two continuous functions is continuous.
6. The value of $H(P)$ remains same after the addition of probability of impossible outcomes.

Since $H(P)$ satisfies all the conditions of valid probabilistic entropy measures so it is a valid probabilistic entropy measure.

Corollary 7: Division of fuzzy entropy function is not a necessary fuzzy entropy.

Proof: Let $F(w)$ & $G(w)$ be two fuzzy entropy functions and $H(w)$ be another function defined in the same domain as $H(w)$ is equal to $F(w)/G(w)$.

Now $H(w)$ is a fuzzy entropy function if it satisfies the condition of valid fuzzy entropy i.e;

1. $H(w)\ is\ equal\ to\ zero$, when w is crisp, since $F(w) = G(w)$ is equal to zero, when w is crisp.
2. Now $H(w)$ is equal to $H(\bar{w})$; $H(w) is\ equal\ to\ F(w)/G(w)$
   This implies $H(\bar{w})$ is equal to $F(\bar{w})/G(\bar{w})$, since $F(w)$ & $G(w)$ are valid entropy so $F(\bar{w})$ is equal to $F(w)$, and $G(\bar{w})$ is equal to $G(w)$, thus $H(\bar{w}) = F(w)/G(w) = H(w)$
3. $H(w)$ attains maximum value if $\mu_w(x) = 0.5\ \forall x \in w$. Since $F(w)$ & $G(w)$ attain maximum value at $\mu_w(x) = 0.5\ \forall x \in w$ due to valid fuzzy entropy hence their division also attains maximum value.
4. Since $H(w) = F(w)/G(w)$
   Then $H(w^*) = \frac{F(w^*)}{G(w^*)}$, As F & G is valid fuzzy entropy so $F(w)$ is greater than and equal to $F(w^*)$ and $G(w)$ is greater than and equal to $G(w^*)$ but it is not necessary that $\frac{F(w)}{G(w)} \geq \frac{F(w^*)}{G(w)}$, therefore it is not necessary $H(w) \geq \frac{F(w^*)}{G(w^*)} = H(w^*)$,
   $H(w)$ is greater than and equal to $H(w^*)$ is only when $F(w) \geq\ G(w) \geq G(w^*) \geq F(w^*)$

Since $H(w)$ does not satisfy all the conditions of valid fuzzy entropy measures it is valid only under certain conditions discussed in the fourth condition.

Remark: The above result is true only when $\frac{0}{0}$ *is considered as* 0.

4. Discussion

To solve different types of mathematical and science problems, differential calculus is used as a significant content. It is concerned with the rate of quantity changes. In differential calculus, the primary objective of study is to find the derivative of functions, concepts of differential and their applications. At a given value the rate of change of any function is illustrated by the differentiation. The best linear approximation of a single variable real-valued function is described by the derivative of a function. In most of sciences, derivative is a significant concept; like in physics derivative is a used concept of rate of change like in motion, velocity and acceleration, law of motion and others, and the derivative concept is used in chemistry as to illustrate the rate of a chemical reaction. In OR, the concept of derivative is utilized to optimize the transport the materials and to design the factories. This concept is also applied to determine maxima and minima of any function.

As information theory is an important topic in recent sciences which is utilized to handle the various practical problems of real life. Generally, in information theory, the studies are related to the storage, communication and quantification of information. Basically, for all types of studies or research, a researcher requires information as data or processed information for further used. Thus, for all types of subject studies researchers use the content of information theory, without information theory it is not easy to complete the studies or research. The application like loss-less data compression, channel coding and lossy data compression are involved in information theory as general topics. The various technical fields like the growth of the internet, the analysis of linguistics and of human perception, the feasibility of mobile phones, the knowledge of black holes and etc, have been critically influenced by information theory.

Along with significant applications, information theory and differential calculus are important concepts of mathematical theory. So, it is a good thought to hybridize the studies of information theory and differential calculus with each other. In information theory, entropy acts as a significant measure (function), which quantifies the uncertainty of a random variable or the outcome of a random process. Thus, here we relate entropy function with differential calculus by using some basic theorems of differential derivative. These theorems can be utilized as a necessary condition to verify the validity of entropy function. Since, it is showed that all entropy functions satisfied RT, thus by using this result the continuity of entropy functions can be easily checked in a given domain. It can be utilized to find the unique maximum value of entropy function and can also be used to show that the maximum point is most uncertain point (0.5). It can also be used to show that the entropy functions are strictly increasing and decreasing in $[0, 0.5]$ and $[0.5, 1]$ respectively. RT can also be used to check that at the boundary points (0 and 1) the value of entropy function is minimum. Since, above, it is also proved that every entropy function neither increases nor decreases in its domain. This statement shows that function is symmetric with respect to most fuzzy (uncertain) point.

LMVT is used to find the point of maxima in entropy function by using formula:

$$f'(c) = \left(\frac{f(b) - f(a)}{(b - a)}\right); c \in (a, b).$$

Here it is also stated that entropy function with respect to another entropy function does not satisfy CMVT and it is also stated under what condition entropy function satisfies CMVT. This statement is used in using L' Hospital rule in "Limit" problems and derivative problems.

The theorems which are discussed in the some properties of the subsection of entropy function' are also useful. These theorems are used to find the new entropy function by the uses of existing entropy function. Thus, new entropy measures may be derived by using the existing measures. By the hybridization of two different entropy functions the limitation of both functions can be removed, and their significance may be increased.

Since, the notion of information theory is included in the concepts of "Mathematics", "Physics", "Computer Science", "Statistics", "Neurobiology" and "Electrical Engineering". Thus, information theory has different applications in various areas, like "Statistical Inference", "Cryptography", "Natural Language", "Thermal Physics", "Processing", "Plagiarism Detection" "Bioinformatics", "Linguistics in Language" and "Pattern Recognition". "Source Coding", "Channel Coding", "Algorithmic Complexity Theory", "Algorithmic Information Theory", "Information Theoretic Security" and "Measures of Information", which are significant subfields of information theory. In these areas differential calculus is also significant, so proposed work is also significant in related areas.

5. Conclusions

In this paper, it is shown that the entropy function satisfies the RT and LMVT. So, by using these properties of entropy function the point of maximum uncertainty can be easily calculated and checked that the most uncertain value has maximum uncertainty. RT and LMVT can be used as a necessary condition of Entropy function. Here it is also stated under what condition entropy function satisfies Cauchy Mean value Theorem. As RT and LMVT are used to solve many mathematical problems or other field's problems, also these theorems can be used in information theory. To increase the database of entropy, some function theorems have been discussed. By applying these theorems new entropy functions can be derived by using the existing entropy functions. This study is an open gate for other studies for related fields and can be used in future works.

Conflict of Interest

The authors declare no conflict of interest.

REFERENCES

[1] J. Willard Gibbs, "Discussion of the thermodynamics analogy", in Elementary principles of statistical mechanics, Ist ed, CSS, march 1902, pp. 193-215.

[2] Fisher R., "On the mathematical foundations of theoretical statistics", Philos. Trans. Roy. Soc. London A., vol. 222, pp. 309-368, 1922.

[3] Shannon C. E., "A mathematical theory of communication", The Bell System Technical Journal, vol. 27, pp. 379-423, 1948.

[4] Kullback S., Leibler R.A., "On information and sufficiency", Ann. Math. Stat., vol. 22, pp. 79-86, 1951.

[5] Renyi A., "On measure of entropy and information", Proceedings of the fourth Berkeley symposium on mathematical statistics and probability, CA: University of California Press, Berkeley, Calif, USA, vol. 1, pp. 547-561, 1961.

[6] Kapur J. N., "Generalization entropy of order α and type β", The mathematical seminar (Delhi), vol. 4, pp. 78-84, 1967.

[7] Tsallis C., "Possible generalization of Boltzmann-Gibbs statistics", Journal of statistical physics, vol. 52, pp. 479-487, 1988.

[8] Sharma O., Tiwari P., Gupta, P., "Fuzzy soft matrices entropy-application in data reduction", fuzzy soft matrices, International Journal of Fuzzy System Application (IJFSA), vol. 7, no. 3, 2018.

[9] Sharma O., and Gupta P., "Probabilistic entropy measures derived by using conic-section equation and their application in dimension reduction", Journal of Statistics and Management Systems, vol. 22, no. 6, 2019. https://doi.org/10.1080/09720510.2019.1596593.

[10] Zadeh L. A., "Fuzzy sets, Information and control", Information and Control, vol. 8, no. 3, pp. 338-353, 1965. https://doi.org/10.1016/S0019-9958(65)90241-X

[11] Zadeh L. A., "Probability measure of fuzzy events", Journal of mathematical analysis and application, vol. 23, pp. 421-427, 1968.

[12] De Luca A., and Termini S., "A definition of a non-probabilistic entropy in the setting of fuzzy entropy", Information and control, vol. 20, pp. 301-312, 1972.

[13] Bhandari D., and Pal N. R., "Some new information measure of fuzzy sets", Information Sciences, vol. 67, pp. 209-228, 1993.

[14] Sharma O., Sonia, Gupta P., "Some Trigonometric Entropy for Fuzzy Rough Set and their Application in Medical Area", International Journal for Research in Engineering Application & Management (IJREAM) (ISSN: 2454-9150), vol. 4, no. 3, pp. 570-578, 2018.

[15] Selin H., "Encyclopedia of the History of Science, Technology and Medicine in Non-Westen Cultures", Springer Dordrecht, 2016, pp. 1-4706.

[16] Rolle, M. Traité d'Algebre. E. Michallet, Paris, 1690.

[17] Rolle, M. Démonstration d'une Méthode pour resoudre les Egalitez de tous les degrez, 1691.

[18] Aczel J., Doroczy Z., "On measure of information and their characterizations", Academic Press, New York, 1975.

[19] Mathai A., Rathie R. N., "Information theory and statistics", Wiley Eastern, New Delhi, 1974.

[20] Cajori, Florian., "A History of Mathematics", TMC, pp. 1-435, 1909.

[21] Li Z., Deng J., Zhow B., Liu Y., Cai Z., "An entropy measurement method of quantum information system under uncertain environment", International Journal of Intelligence Science, vol.8, no. 2, pp. 29-41, 2018. https://doi.org/10.4236/ijis.2018.82002.

[22] Robinson W.D., "Entropy and Uncertainty", vol. 10, pp. 493-509, 2008, https://doi.org/10.3390/e10040493.

[23] Zhang Y., Huang F., Deng X., Jing W., "A new total uncertainty measure from a perspective of maximum entropy requirement", Entropy, vol. 23, no. 8, 1061, 2021. https://doi.org/10.3390/e23081061.

[24] Namdari A., Li(Steven) Z., "A review of entropy measures for uncertainty qualification of stochastic processes", Advances in Mechanical Engineering, vol. 11, no. 6, pp. 1-14, 2019. https://doi.org/10.1177/1687814019857350.

[25] Chakraborty S., Paul D., Das S., "A new measure of uncertainty with some application", IEEE Symposium on Information Theory (ISIT), pp. 1-23, 2021.arXiv:2105.00316v2 [cs.IT] 5 May 2021.s