STATISTICS, OPTIMIZATION AND INFORMATION COMPUTING
Stat., Optim. Inf. Comput., Vol. **x**, Month 202x, pp 0–16.
Published online in International Academic Press (www.IAPress.org)

Contingency Uniformity Measure: An approach for spread characterization in Contingency tables

Ratnesh Kumar Singh, Naveen Kumar, Vivek Vijay*

Department of Mathematics, Indian Institute of Technology Jodhpur, 342030, Rajasthan, India

Abstract This paper introduces the Contingency Uniformity Measure (CUM), a normalized entropy that scales Shannon entropy to the range $[0, 1]$, enabling fair comparisons across contingency tables of different dimensions. CUM retains key properties such as non-negativity, reaches its maximum at the uniform distribution, and satisfies weighted additivity. We formulate and solve three optimization problems, using CUM, under realistic constraints, fixed marginal distributions, a cost matrix, and cost variance. This is demonstrated through a real dataset of cost matrix obtained using distance matrix. Our results show that CUM is an effective, standardized measure for analyzing uncertainty and supporting decision-making in diverse, constraint-driven systems.

Keywords Entropy, Contingency Table, Maximum Entropy Principle, Probability distribution

AMS 2010 subject classifications 94A17,62H17,62B10

DOI: 10.19139/soic-2310-5070-2878

1. Introduction

Entropy is a central concept in both theoretical and applied sciences, serving as a measure of uncertainty, disorder, or information content in a system. Originally introduced in the field of thermodynamics by Rudolf Clausius in the 19th century, the concept was later reinterpreted in a probabilistic context by Claude Shannon in his foundational work on information theory [1]. Shannon entropy quantifies the average information produced by a random variable. Entropy plays a key role in areas such as machine learning, communication, healthcare, economics and finance, physics, network theory, cryptography, and statistical mechanics making informed decisions due to its ability to quantify uncertainty and information content [2, 3, 4, 5, 6, 7, 8, 9, 10, 11, 12, 13, 14, 15, 16, 17, 18, 19, 20, 21].

A contingency table is a statistical tool that is used to display the frequency distribution of categorical variables to analyze the relationship. It is utilized in various fields, such as medicine, social sciences, market behavior, and machine learning, for studies like classification and association analysis [22]. The principle of maximum entropy is a modeling framework to derive the least-biased null distribution under marginal constraints, enabling the detection of higher-order interactions in multi-dimensional contingency tables [23]. The maximum entropy principle is applied to derive a probability distribution for categorical data that adheres to specified confidence interval bounds [24]. Also, the principle of maximum entropy, under some marginal constraints, provides estimates of missing data in a contingency table [25]. Gzyl [26] provides a solution for deriving joint distributions from known marginals ensuring the most unbiased estimation consistent with the given marginals. Furthermore, the study [27] applies the maximum entropy principle to derive joint probability distributions from known marginals

*Correspondence to: Vivek Vijay (Email: vivek@iitj.ac.in). Department of Mathematics, Institute of Technology Jodhpur, 342030, Rajasthan, India.

ISSN 2310-5070 (online) ISSN 2311-004X (print)

and expected values, accommodating constraints such as probability bounds, with applications in areas such as transportation modeling. The maximum value of entropy depends on the number of possible outcomes of the variable, which poses a challenge when comparing entropy values across different multidimensional contingency tables. The COOLCAT algorithm clusters categorical data using entropy to reduce uncertainty, preserving the structure of the data while minimizing its complexity [28]. Hong and Kim [29] examine how mutual information and redundancy, based on entropy, can be used to analyze relationships in multi-dimensional tables without relying on the assumption of a specific distribution. In their work [30], Patrick and Sarah employ maximum entropy and minimum norm methods to estimate multivariate probabilities from limited categorical data using equality and inequality constraints. An entropy-based normalized mutual information measure is presented in [31] to assess dependency in a two-way contingency table.

The motivation of this study arises from the challenge of quantitatively assessing the degree of spread or concentration in categorical data represented through contingency tables, where conventional entropy measures such as Shannon entropy fail to provide meaningful comparisons across tables of different dimensions. To address this limitation, the article introduces the Contingency Uniformity Measure (CUM), a normalized entropy-based index bounded between 0 and 1, which quantifies how uniformly data are distributed across categories, thereby allowing for consistent comparison among contingency tables of varying sizes. The proposed measure achieves values close to 1 for highly uniform distributions and near 0 for highly concentrated ones. Building on this formulation, the article develops optimization models to maximize or minimize CUM under various constraints, namely fixed marginals, linear cost constraints, and quadratic cost variance constraints demonstrating its applicability in domains such as logistics, healthcare and financial risk management. To contextualize CUM within broader measures of diversity and uncertainty, a comparative analysis with the Gini–Simpson Index and Theil's Uncertainty Coefficient is presented, highlighting CUM's superior interpretability and consistency with information theoretic principles. Furthermore, the study illustrates computational practicality through a real-world optimization example using a 10×10 city distance matrix and empirically validates the approach using State/UT-wise data on public and private hospitals enrolled under the Ayushman Bharat–Pradhan Mantri Jan Arogya Yojana (AB-PMJAY) scheme from $2018-19$ to $2024-25$. A one-tailed two-sample t test at the 5% significance level confirms that private hospitals exhibit significantly higher uniformity in distribution, reinforcing the analytical utility and interpretive strength of the proposed CUM framework.

The article is organized as follows. In Section 2, we present some definitions required in subsequent sections. In Section 3, we define the contingency uniformity measure and discuss its properties, along with a comparison to other related measures. We present the optimization problem for the contingency uniformity measure in Section 4. Some real life applications are discussed in Section 5. The article is concluded with some discussions in Section 6.

2. Preliminaries

Let Ω be a sample space, and $A \subseteq \Omega$ be an event with its probability $P(A)$. The information associated with event A is defined mathematically based on its probability, rather than the content of the event itself. Intuitively, the less likely an event is, the more informative or surprising it is considered to be. For events, where $P(A) \neq 0$, the information content of A is given by

$$I(A) = \ln\left(\frac{1}{P(A)}\right) = -\ln P(A).$$

Definition 1. *Let $\Omega = \{\omega_1, \omega_2, \ldots, \omega_n\}$ be a discrete sample space, and $P(\omega_i)$ be the probability of the event ω_i. The uncertainty associated with the sample space is quantified by the average information content, where the information of each event is defined as $I(\omega_i) = -\ln P(\omega_i)$. This average gives a measure of the overall uncertainty or entropy of the system (Shannon entropy).*

$$H(X) = \sum_{i=1}^{n} P(\omega_i) I(\omega_i) = -\sum_{i=1}^{n} P(\omega_i) \ln P(\omega_i), \tag{1}$$

for more details, see [32, 33].

One important application of entropy is the principle of maximum entropy, useful in obtaining the probability distribution under certain constraint, formulated as the following optimization problem:

$$\begin{aligned}
&\text{Maximize:} \quad && H(X) = -\sum_{i=1}^{n} P(w_i)\log P(w_i) \\
&\text{Subject to} \quad && \sum_{i=1}^{n} P(w_i) = 1 \quad \text{(normalization)}, \\
& && \sum_{i=1}^{n} P(w_i)F(w_i) = \mu \quad \text{(moment constraint)}.
\end{aligned}$$

It estimates the probability distribution for a system under moment constraints and remains valid in highly uncertain scenarios. For more details, see [34].
In the following section, we define the contingency uniformity measure (CUM) and discuss its properties.

3. Contingency Uniformity Measure(CUM)

A contingency table with a large number of cells can have higher entropy than a table with fewer cells, even if both exhibit similar patterns of uncertainty. The proposed normalized entropy, CUM, allows meaningful comparison by removing this dependence. It is a measure of uncertainty that scales Shannon entropy relative to its maximum possible values, ensuring a consistent and interpretable measure of randomness across various multidimensional contingency tables.

Definition 2 (**Contingency Uniformity Measure**). *Let $X = \{x_1, x_2, \ldots, x_n\}$ be a discrete random variable and $P = \{p_1, p_2, \ldots, p_n\}$ be the corresponding probability distribution. Let $H_{max}(X) = \max_P H(X)$. A contingency uniformity measure is defined as*

$$H_{CUM} = \frac{H(X)}{H_{max}(X)}. \tag{2}$$

Note that, for a discrete random variable, with n elements in support, $H_{max} = \ln n$. Therefore,

$$H_{CUM} = \frac{H(X)}{\ln n}.$$

We now state some properties of H_{CUM}.

Theorem 1. *For a discrete random variable with support $\{x_1, x_2, \ldots, x_n\}$,*

(i) H_{CUM} *is nonnegative.*

(ii) H_{CUM} *is non-expansible.*

(iii) $0 \leq H_{CUM} \leq 1$.

(iv) H_{CUM} *attains its maximum value at the uniform distribution.*

(v) H_{CUM} *is non-decreasing function of uncertainty.*

(vi) H_{CUM} *does not satisfy additivity property, that is, for any two independent random variables X and Y, $H_{CUM}(X,Y) \neq H_{CUM}(X) + H_{CUM}(Y)$.*

(vii) H_{CUM} *is invariant under probability rescaling.*

(viii) H_{CUM} *is independent of category count.*

(ix) H_{CUM} *is a concave function of X.*

Proof
(i) $H(X)$ is non-negative and $\ln(n) > 0$ implies

$$H_{CUM} = \frac{H(X)}{\ln n} \geq 0.$$

(ii) For any distribution $\{p_1, p_2, \dots p_n\}$, we know that $H(p_1, p_2, \dots p_n, 0) = H(p_1, p_2, \dots p_n)$.
Thus,

$$\frac{H(p_1, p_2, \dots p_n, 0)}{\ln(n+1)} = \frac{H(p_1, p_2, \dots p_n)}{\ln(n+1)} = \frac{H(X)}{\ln n}\frac{\ln n}{\ln(n+1)} = \frac{\ln n}{\ln(n+1)}H_{CUM}.$$

(iii) Since $0 \leq H(X) \leq H_{max}(X)$, we have

$$0 \leq H_{CUM} \leq 1.$$

(iv) As the Shannon entropy is maximum at uniform distribution, we have $H(X) = H_{max}(X) = \ln n$.
Therefore,

$$H_{CUM} = \frac{H(X)}{\ln n} = 1.$$

(v) If the probability distribution becomes more concentrated, the entropy decreases and so does $H_{CUM}(X)$. If the probability distribution becomes more uniform, entropy increases, and so does $H_{CUM}(X)$. Thus, for any two distributions $P = \{p_1, p_2, \dots, p_n\}$ and $Q = \{q_1, q_2, \dots, q_n\}$, if P is more uniform than Q, then $H_{CUM}(P) \geq H_{CUM}(Q)$. That is, $H_{CUM}(X)$ increases monotonically with increasing uniformity of the probability distribution.
(vi) For any two independent random variables X and Y, we have

$$H(X, Y) = H(X) + H(Y). \tag{3}$$

Now

$$H_{max}(X, Y) = \ln(n_X n_Y) = \ln(n_X) + \ln(n_Y). \tag{4}$$

Therefore,

$$\begin{aligned} H_{CUM}(X, Y) &= \frac{H(X, Y)}{H_{max}(X, Y)}, \\ &= \frac{H(X) + H(Y)}{\ln(n_X) + \ln(n_Y)}, \\ &= H_{CUM}(X)\frac{\ln(n_X)}{\ln(n_X) + \ln(n_Y)} + H_{CUM}(Y)\frac{\ln(n_Y)}{\ln(n_X) + \ln(n_Y)}. \end{aligned}$$

This implies that $H_{CUM}(X)$ satisfies a weighted additivity property, where the weights depend on the logarithm of the number of states of each independent variable.

(vii) A function is invariant under rescaling if multiplying the probability by a constant factor does not change its value. If we rescale the probabilities by a factor $c > 0$, $p'_i = cp_i$, we check wether $H_{CUM}(X)$ remains unchanged. The total probability must sum to 1

$$\sum p'_i = 1 \implies c\sum p_i = 1$$

which means $c = 1$, meaning valid probability distribution does not allow arbitrary rescaling. Thus, Shannon Entropy is not affected by simple rescaling unless we allow normalization of the probability afterward. Hence,

H_{CUM} is invariant under probability rescaling.

(viii) The entropy $H(X)$ depends on the number of counts of the categories as $\log|X|$ increases with increasing number of categories. However, by normalizing, we scale the entropy to always fall between 0 and 1, making it independent of the categories counts.

(ix) For two distributions, $p = (p_1, \ldots, p_n)$ and $q = (q_1, \ldots, q_n)$ we have

$$H(\lambda p + (1-\lambda)q) \geq \lambda H(p) + (1-\lambda)H(q), \quad 0 \leq \lambda \leq 1.$$

Divide both side by $\ln n > 0$, we get $H_{CUM}(X)$ to be a concave function. $\square$

Example 1. *Consider following 2×2 contingency tables with different probability distributions:*

100	*0*
0	*0*

Table 1. $H_{CUM} = 0$

67	*33*
0	*0*

Table 2. $H_{CUM} = 0.4591$

50	*50*
0	*0*

Table 3. $H_{CUM} = 0.4999$

33	*33*
0	*34*

Table 4. $H_{CUM} = 0.7924$

15	*25*
25	*35*

Table 5. $H_{CUM} = 0.9703$

25	*25*
25	*25*

Table 6. $H_{CUM} = 1$

As the probability distribution becomes more dispersed across the table, the value of H_{CUM} increases. When the probability mass is concentrated in a single cell, $H_{CUM} = 0$, indicating complete predictability. As the distribution becomes more randomized, H_{CUM} increases, reflecting greater uncertainty. The maximum value of H_{CUM} is attained when the distribution is uniform, representing the highest level of unpredictability.

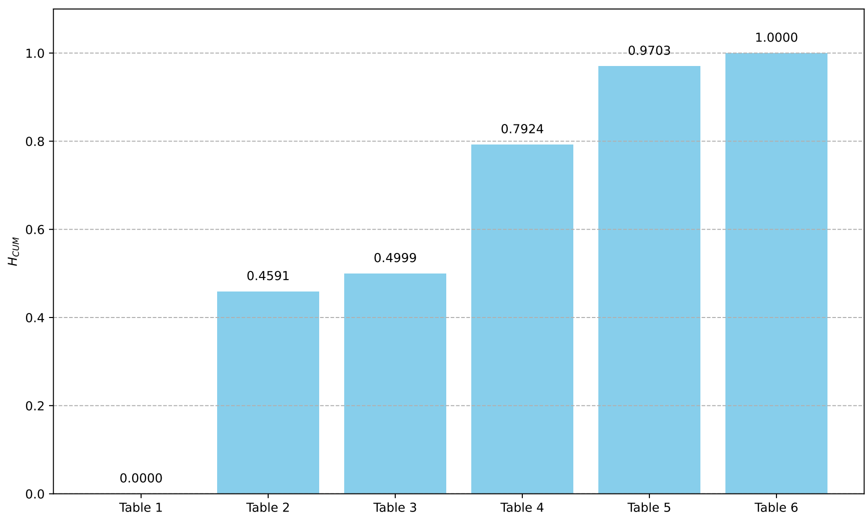

Figure 1. Plot of CUM values for contingency Table $1-6$

Figure (1) clearly shows that H_{CUM} is sensitive to the shape of the distribution.

Table	H	H_{CUM}	GS	U
Table 7	0	0	0	-
Table 8	0.9	0.6493	0.5313	0.1031
Table 9	1.0735	0.7744	0.5625	0.0909
Table 10	1.2555	0.9056	0.6875	0.0002
Table 11	1.3208	0.9528	0.7187	0.0448
Table 12	1.3863	1	0.75	0

Table 13. Comparison between H, H_{CUM}, GS and U

3.1. Comparison between Entropy, Contingency Uniformity Measure, Gini-Simpson Index and Theil's Uncertainty coefficient

Gini-Simpson Index(GS): The Gini–Simpson Index is a statistical measure used to represent inequality or diversity within a distribution. It quantifies the probability that two elements randomly selected from a dataset belong to different categories, or equivalently, the likelihood that a randomly chosen element would be incorrectly classified if assigned according to the class proportions. It is mathematically defined as

$$G = 1 - \sum_{i=1}^{n} p_i^2$$

The Gini-Simpson Index ranges between 0 and 1, where values close to 0 indicate low diversity or high concentration, whereas values near 1 indicate high diversity or uniformity across categories. for more details, see [35].

Theil's uncertainty coefficient(U): The Theil's uncertainty coefficient is a statistical measure used to quantify the degree of association or predictive power between two variables. In the context of information theory and machine learning, it measures how much uncertainty in one variable Y can be explained or reduced by knowing another variable X. Mathematically, it is expressed as

$$U(Y|X) = \frac{H(Y) - H(Y|X)}{H(Y)}$$

Theil's Uncertainty Coefficient ranges between 0 and 1, where values close to 0 indicate no association or complete uncertainty, while values approaching 1 represent perfect association or complete certainty. for more details, see [32]

Example 2. *Consider following 2×2 contingency tables with different probability distributions:*

8	*0*
0	*0*

Table 7

0	*1*
2	*5*

Table 8

1	*1*
5	*1*

Table 9

1	*1*
3	*3*

Table 10

1	*3*
2	*2*

Table 11

2	*2*
2	*2*

Table 12

From Table (13), it can be observed that as entropy increases, both the H_{CUM} and the Gini-Simpson index exhibit a consistent upward trend. In contrast, Theil's uncertainty coefficient does not show a regular or monotonic increase corresponding to changes in entropy. This irregular behavior suggests that Theil's uncertainty coefficient *may not be a reliable criterion to compare different probability distributions, particularly when assessing their degree of randomness or inequality. In terms of interpretation,* H_{CUM} *directly represents the proportion of the maximum possible uncertainty in the joint distribution. It provides a clear probabilistic meaning, values near* 1 *indicate high randomness and diversity, while values near* 0 *indicate complete concentration or certainty. The Gini–Simpson Index, quantify concentration and diversity within a distribution, respectively; however, they do not possess a direct connection to uncertainty in the information-theoretic sense. While these indices effectively capture dominance or evenness among categories, they lack the explicit probabilistic interpretation that entropy-based measures provide. From a theoretical foundation perspective,* H_{CUM} *is grounded in information theory, derived from Shannon's entropy, and satisfies key properties such as non-negativity and weighted additivity. In contrast, the Gini–Simpson index originates from ecological and probabilistic diversity studies, serving as practical indicators of heterogeneity or inequality but without the same foundational link to information content. In general* H_{CUM} *offers clearer probabilistic interpretation and compatibility with information-theoretic analysis and optimization frameworks, making it preferable when modeling uncertainty or diversity in categorical systems.*

4. Optimization problems in contingency table

Contingency tables are widely used in statistics, machine learning, and information theory to analyze relationships among categorical variables. A contingency table represents the joint distribution of two or more categorical variables. Optimization problems related to contingency tables often aim to adjust cell values to satisfy constraints while maximizing or minimizing the underlying objective function.
Here, we maximize or minimize the $H_{CUM}(X, Y)$, depending on the problem, with some probability constraints.

Problem 1. *Let T be a* $m \times n$ *contingency table, where each cell* (p_{ij}) *represents the probability of the joint occurrence of categories* i *of variable* X *and category* j *of variable* Y*. The objective is to maximize(or minimize) contingency uniformity measure subject to certain constraints imposed by the marginal distributions. That is,*

$$\max_{T} \; H_{CUM}(X, Y)$$

subject to

$$\sum_i \sum_j p_{ij} = 1, \;\; \forall i, j$$

$$\sum_j p_{ij} = p_{i+}, \;\; \forall j$$

$$p_{ij} \geq 0, \;\; \forall i, j$$

Solution 1. *Define the Lagrangian*

$$L = -\sum_i^m \sum_j^n p_{ij} \frac{\ln p_{ij}}{\ln(mn)} + \lambda \left(\sum_i^m \sum_j^n p_{ij} - 1 \right) + \sum_i^m \mu_i \left(\sum_j^n p_{ij} - p_{i+} \right).$$

Differentiating the Lagrangian with respect to p_{ij}*, we get*

$$-\frac{1}{\ln(mn)} (1 + \ln(p_{ij})) + \lambda + \mu_i = 0.$$

Solving for p_{ij}*, we get*

$$p_{ij} = e^{(\lambda + \mu_i)\ln(mn) - 1}. \tag{5}$$

Using the constraints, we obtain the coefficients satisfying

$$\sum_{j=1}^{n} e^{(\lambda+\mu_i)\ln(mn)-1} = p_{i+},$$

this implies,

$$e^{(\lambda+\mu_i)\ln(mn)-1} = \frac{p_{i+}}{n} \tag{6}$$

from equations (5) and (6), we get

$$p_{ij} = \frac{p_{i+}}{n},\ \forall j = 1, 2, \ldots, n.$$

The following example illustrates the application of the above theorem in a practical marketing context, showing how the concept of maximizing H_{CUM} can be used to model the most uncertain customer behavior.

Example 3. *A company runs a marketing campaign for Product A. Customers are segmented into three age groups: Young, Adult, and Senior. The company observes how customers from different age groups respond to the product (that is, Buy or Not Buy), but only the age group distribution of its customers is known (that is,Young = 20%, Adult = 60%, Senior = 20%). Through an optimization problem, the company aims to maximize H_{CUM} to model the most uncertain scenario to understand the behavior of the customers. This allows the company to model the case where customer responses are completely unpredictable, providing a benchmark for comparison against actual observed behavior. The contingency table is defined as:*

Age Group ↓ / Response →	Buy	Not Buy	Row Total
Young	p_{11}	p_{12}	0.2
Adult	p_{21}	p_{22}	0.6
Senior	p_{31}	p_{32}	0.2

Table 14. Initial data for Product A

To achieve maximum H_{CUM}, the joint distribution is:

$$p_{11} = p_{12} = 0.1, \quad p_{21} = p_{22} = 0.3, \quad p_{31} = p_{32} = 0.1$$

Age Group ↓ / Response →	Buy	Not Buy	Row Total
Young	0.1	0.1	0.2
Adult	0.3	0.3	0.6
Senior	0.1	0.1	0.2

Table 15. Joint distribution for Product A maximizing H_{CUM}

That is, each age group is evenly split between buying and not buying.

Note that obtaining the joint probability distribution in a contingency table based on given marginal probabilities results in a uniform distribution along the corresponding rows or columns, and maximizing CUM yields a consistent value for the marginals, making it independent of the size of the table. In the next problem, we present a cost optimization framework under highly uncertain scenarios, which is particularly useful in manufacturing industries to obtain reliable estimates.

Problem 2. *Let T be a $m \times n$ contingency table, where each cell (p_{ij}) represents the probability of the joint occurrence of categories i of variable X and category j of variable Y. Given the cost matrix $[R_{ij}]$, associated with the joint occurrence of events $X = i$ and $Y = j$, the objective is to determine the joint probability distribution p_{ij}*

that either maximizes or minimizes the contingency uniformity measure, subject to constraint that the expected cost calculated as the weighted sum of R_{ij} with respect to p_{ij}, is equal to a fixed constant R_0. That is,

$$\max_{T} \ H_{CUM}(X,Y)$$

subject to

$$\sum_{i}\sum_{j} p_{ij} = 1, \ \ \forall i,j \tag{7}$$

$$\sum_{i}\sum_{j} R_{ij}p_{ij} = R_0 \tag{8}$$

$$p_{ij} \geq 0, \ \ \forall i,j.$$

Solution 2. *The Lagrangian function is*

$$L = -\sum_{i}^{m}\sum_{j}^{n} p_{ij}\frac{\ln p_{ij}}{\ln(mn)} + \lambda\left(\sum_{i}^{m}\sum_{j}^{n} p_{ij} - 1\right) + \mu\left(\sum_{i}^{m}\sum_{j}^{n} R_{ij}p_{ij} - R_0\right).$$

Differentiating the Lagrangian with respect to p_{ij}, we get

$$-\frac{1}{\ln(mn)}(1+\ln p_{ij}) + \lambda + \mu R_{ij} = 0.$$

Solving for p_{ij}, we get

$$p_{ij} = e^{(\lambda+\mu R_{ij})\ln(mn)-1}. \tag{9}$$

Using the equation (7), we obtain the coefficients satisfying

$$e^{\lambda\ln(mn)-1}\sum_{i}^{m}\sum_{j}^{n} e^{\mu R_{ij}\ln(mn)} = 1,$$

this implies,

$$e^{\lambda\ln(mn)} = \frac{e}{\sum_{i}^{m}\sum_{j}^{n} e^{\mu R_{ij}\ln(mn)}}. \tag{10}$$

From equations (9) and (10), we get

$$p_{ij} = \frac{e^{\mu R_{ij}}}{\sum_{i}^{m}\sum_{j}^{n} e^{\mu R_{ij}}}. \tag{11}$$

Putting the value of p_{ij} in equation (8), we get

$$\sum_{i}^{m}\sum_{j}^{n} R_{ij}\frac{e^{\mu R_{ij}}}{\sum_{i}^{m}\sum_{j}^{n} e^{\mu R_{ij}}} = R_0,$$

this implies,

$$\sum_{i}^{m}\sum_{j}^{n} (R_{ij} - R_0)e^{\mu R_{ij}} = 0. \tag{12}$$

The nonlinear equation (12) can be numerically solved for μ using iterative root-finding methods such as the Newton–Raphson, Secant, or Bisection algorithms. Once the optimal value of μ is obtained, it can be substituted *into equation (11) to compute the desired result. Each iteration requires evaluating two double summations over $m \times n$ elements, resulting in a computational complexity of $O(mn)$ per iteration. Consequently, the overall computational time increases linearly with the grid size. A weighted average can never exceed the largest value or fall below the smallest value of the set being averaged. If $R_0 > R_{max}$ or $R_0 < R_{min}$ the equation (12) will have no real solution for μ.*

The following example demonstrates the application of the above theorem to a logistics optimization scenario, where the goal is to maximize H_{CUM} under cost constraints.

Example 4. *A logistics company aims at optimizing the uncertainty of shipments between warehouses (X) and stores (Y) while ensuring that the total expected cost does not exceed a specified budget R_0. Assume there are three warehouses W_1, W_2, W_3 and three stores S_1, S_2, S_3. Let p_{ij} represents the proportion of total shipments from warehouse x_i to store y_j, and let R_{ij} denote the cost of shipping from warehouse i to store j. The objective is to maximize the contingency uniformity measure $H_{CUM}(X, Y)$, which corresponds to identifying the shipment pattern that exhibits the highest level of uncertainty or randomness in allocation (that is, shipments are spread out with minimal identifiable preference between specific warehouses and stores), subject to the expected cost constraint.*

The shipment proportions are defined as:

Warehouse $X \downarrow$ / Store $Y \rightarrow$	y_1	y_2	y_3
x_1	p_{11}	p_{12}	p_{13}
x_2	p_{21}	p_{22}	p_{23}
x_3	p_{31}	p_{32}	p_{33}

Table 16. Proportion of total shipments from x_i to y_j

The associated cost matrix is:

Warehouse $X \downarrow$ / Store $Y \rightarrow$	S_1	S_2	S_3
W_1	4	6	8
W_2	5	3	7
W_3	2	1	3

Table 17. Cost matrix for shipping from warehouses to stores

Assume the total expected cost budget is $R_0 = 6$. Then, using equation (12), the computed value of $\mu = 0.435$, and by substituting it into equation (11), the optimized shipment proportions for maximum H_{CUM} are given in the following table 18

Warehouse $X \downarrow$ / Store $Y \rightarrow$	y_1	y_2	y_3
x_1	0.0613	0.1463	0.3497
x_2	0.0948	0.0397	0.2261
x_3	0.0257	0.0167	0.0397

Table 18. Proportion of total shipments from x_i to y_j for maximum H_{CUM}

In the following example, we use a cost matrix obtained using a distance matrix of 20 different cities of a state of Uttar Pradesh, India.

Example 5. *The Government of Uttar Pradesh plans to operate intercity bus services connecting 20 selected cities. For planning and data collection, the pairwise distances between city centers are available. According to the Ministry of Housing and Urban Affairs, the operating cost has been estimated at 99.2 per kilometer [36] (page*

103). Let the 20 cities be denoted by $C_1, C_2, \ldots, C_{20}$. The flow of bus trips from city C_i to city C_j is represented by the probability p_{ij}, indicating the proportion of total bus trips allocated from C_i to C_j. The objective is to determine the probability matrix $P = [p_{ij}]$ that maximizes the evenness (or uncertainty) of these flows while satisfying a cost constraint derived from the distance matrix. The distance matrix is denoted by $D = [d_{ij}]$, where d_{ij} is the measured road distance (in kilometers) between cities C_i and C_j, as given in table 19. The corresponding cost matrix is $R = [R_{ij}]$, computed as $R_{ij} = c \times d_{ij}$, where c represents the cost per kilometer. The optimization seeks to maximize the Contingency Uniformity Measure (H_{CUM}), which reflects the evenness of bus service allocation across all city pairs, subject to an average cost constraint.
Assume the total expected cost budget is $R_0 = 80000$. Then, using equation (12), the computed value of $\mu = 0.000317106$, and by substituting it into equation (11), the optimized shipment proportions for maximum H_{CUM} are shown in the following table 20.

This formulation provides the most unbiased probability distribution consistent with known cost constraints, enabling informed decision-making under uncertainty.
The next problem derives the probability distribution under a constraint on the variation in the contingency table, measured using the data variance.

Problem 3. *Let T be a $m \times n$ contingency table, where each cell (p_{ij}) represents the probability of the joint occurrence of categories i of variable X and category j of variable Y. Given a cost matrix $[R_{ij}]$, representing the cost associated with the joint occurrence of events $X = i$ and $Y = j$, the objective is to determine the joint probability distribution p_{ij} that either maximizes or minimizes the contingency uniformity measure, subject to constraint that the weighted sum of the squared deviations of the cost values R_{ij} from their mean R_{mean}, using p_{ij} as weights, is equal to a fixed constant R_1. That is,*

$$\max_{T} \; H_{CUM}(X, Y)$$

subject to

$$\sum_i \sum_j p_{ij} = 1, \;\; \forall i, j \tag{13}$$

$$\sum_i \sum_j (R_{ij} - R_{mean})^2 p_{ij} = R_1 \tag{14}$$

$$p_{ij} \geq 0, \;\; \forall i, j.$$

Solution 3. *The Lagrangian function is*

$$L = -\sum_i^m \sum_j^n p_{ij} \frac{\ln p_{ij}}{\ln(mn)} + \lambda \left(\sum_i^m \sum_j^n p_{ij} - 1 \right) + \mu \left(\sum_i^m \sum_j^n (R_{ij} - R_{mean})^2 p_{ij} - R_1 \right).$$

Differentiating the Lagrangian with respect to p_{ij}, we get

$$-\frac{1}{\ln(mn)}(1 + \ln p_{ij}) + \lambda + \mu (R_{ij} - R_{mean})^2 = 0,$$

solving for p_{ij}, we get

$$p_{ij} = e^{(\lambda + \mu (R_{ij} - R_{mean})^2) \ln(mn) - 1}. \tag{15}$$

Using the equation (13), we obtain the coefficients satisfying

$$e^{\lambda \ln(mn) - 1} \sum_i^m \sum_j^n e^{\mu (R_{ij} - R_{mean})^2 \ln(mn)} = 1,$$

Table 19. Distance matrix $D = [d_{ij}]$ (in km) between 20 cities in Uttar Pradesh

	Noida	Moradabad	Mathura	Saharanpur	Ayodhya	Jhansi	Etawah	Muzaffarnagar	Sitapur	Firozabad
Lucknow	413.4	315.0	340.1	470.4	127.6	296.2	190.7	442.5	93.8	301.5
Kanpur	379.4	288.5	309.8	447.1	167.4	222.1	117.7	419.0	130.7	251.9
Varanasi	671.9	565.7	628.3	756.4	178.4	343.9	289.8	730.0	276.2	437.3
Agra	178.5	219.1	54.6	256.5	489.8	235.9	188.4	233.9	361.2	43.3
Meerut	62.6	98.8	147.5	91.1	576.1	375.8	297.8	49.2	382.2	178.8
Prayagraj	603.2	498.6	561.7	692.3	155.9	319.1	265.8	666.2	263.8	413.3
Bareilly	217.4	107.1	228.8	231.1	328.0	450.7	348.8	191.4	188.3	288.5
Aligarh	142.4	156.7	101.8	197.4	446.4	318.3	245.4	173.7	303.1	83.5
Ghaziabad	31.2	125.4	134.8	102.8	557.9	355.5	278.1	75.3	369.8	163.7
Gorakhpur	736.5	627.4	700.2	827.7	157.3	428.1	344.7	802.3	251.1	513.2

Table 20. Probability matrix $P = [p_{ij}]$ for 20 cities in Uttar Pradesh

	Noida	Moradabad	Mathura	Saharanpur	Ayodhya	Jhansi	Etawah	Muzaffarnagar	Sitapur	Firozabad
Lucknow	1.28E-06	5.79E-08	1.27E-07	7.68E-06	1.59E-10	3.20E-08	1.16E-09	3.19E-06	5.50E-11	3.79E-08
Kanpur	4.39E-07	2.51E-08	4.91E-08	3.69E-06	5.57E-10	3.11E-09	1.17E-10	1.53E-06	1.76E-10	7.95E-09
Varanasi	0.00435	0.000154	0.00110	0.06205	7.88E-10	1.44E-07	2.62E-08	0.02705	1.71E-08	2.71E-06
Agra	7.90E-10	2.83E-09	1.60E-11	9.19E-09	1.41E-05	4.81E-09	1.08E-09	4.51E-09	2.48E-07	1.12E-11
Meerut	2.06E-11	6.44E-11	2.98E-10	5.05E-11	2.14E-04	3.92E-07	3.37E-08	1.35E-11	4.79E-07	7.98E-10
Prayagraj	0.000501	1.87E-05	0.000136	0.00826	3.88E-10	6.58E-08	1.23E-08	0.00363	1.16E-08	1.27E-06
Bareilly	2.69E-09	8.36E-11	3.85E-09	4.13E-09	8.71E-08	4.13E-06	1.68E-07	1.19E-09	1.08E-09	2.51E-08
Aligarh	2.54E-10	3.98E-10	7.08E-11	1.43E-09	3.61E-06	6.42E-08	6.48E-09	6.79E-10	3.98E-08	3.98E-11
Ghaziabad	7.68E-12	1.49E-10	2.00E-10	7.30E-11	1.20E-04	2.07E-07	1.81E-08	3.08E-11	3.24E-07	4.96E-10
Gorakhpur	0.03318	0.00107	0.01059	0.58456	4.06E-10	2.03E-06	1.47E-07	0.26292	7.75E-09	2.95E-05

this implies,

$$e^{\lambda \ln(mn)} = \frac{e}{\sum_i^m \sum_j^n e^{\mu(R_{ij}-R_{mean})^2 \ln(mn)}}. \tag{16}$$

From equations (15) and (16), we get

$$p_{ij} = \frac{e^{\mu(R_{ij}-R_{mean})^2}}{\sum_i^m \sum_j^n e^{\mu(R_{ij}-R_{mean})^2}}. \tag{17}$$

Putting the value of p_{ij} in equation (14), we get

$$\sum_i^m \sum_j^n (R_{ij} - R_{mean})^2 \frac{e^{\mu(R_{ij}-R_{mean})^2}}{\sum_i^m \sum_j^n e^{\mu(R_{ij}-R_{mean})^2}} = R_1,$$

this implies,

$$\sum_i^m \sum_j^n ((R_{ij} - R_{mean})^2 - R_1) e^{\mu(R_{ij}-R_{mean})^2} = 0. \tag{18}$$

The nonlinear equation (18) can be numerically solved for μ using iterative root-finding methods such as the Newton–Raphson, Secant, or Bisection algorithms. Once the optimal value of μ is obtained, it can be substituted into equation (17) to compute the desired result. Each iteration requires evaluating two double summations over $m \times n$ elements, resulting in a computational complexity of $O(mn)$ per iteration. Consequently, the overall computational time increases linearly with the grid size. Because of increased nonlinearity, convergence may be slightly slower.

The following example illustrates how the above theorem can be applied to an employee–task assignment problem, where the objective is to maximize H_{CUM} while balancing cost efficiency and fairness.

Example 6. *A company aims to assign employees (X) to tasks (Y) such that the assignment is both cost-effective and fair. Each assignment incurs a cost R_{ij}, which reflects the suitability of employee i for task j; a higher cost indicates a greater mismatch. In addition to minimizing cost, the company seeks to balance the objectives: fairness (modeled by maximizing the contingency uniformity measure H_{CUM}). To avoid scenarios where only the easiest or hardest tasks are allocated (that is, extreme assignments), the company imposes a constraint on the variance of assignment costs. This ensures a fair distribution of workload across employees while maintaining efficiency.*
Assume there are three employees and three tasks. Let p_{ij} represent the probability of assigning employee i to task j. The initial probability matrix is given in Table 21, and the cost matrix is presented in Table 22.

Employee $\downarrow$ / Task $\rightarrow$	y_1	y_2	y_3
x_1	p_{11}	p_{12}	p_{13}
x_2	p_{21}	p_{22}	p_{23}
x_3	p_{31}	p_{32}	p_{33}

Table 21. Initial assignment probabilities p_{ij}

Employee $\downarrow$ / Task $\rightarrow$	y_1	y_2	y_3
x_1	4	6	8
x_2	5	3	7
x_3	2	1	3

Table 22. Assignment cost matrix R_{ij}

Employee ↓ / Task →	y_1	y_2	y_3
x_1	0.0847	0.0969	0.1633
x_2	0.0862	0.0924	0.1203
x_3	0.1112	0.1524	0.0924

Table 23. Optimal assignment probabilities under cost variance constraint

The variance constraint on the assignment cost is fixed at $R_1 = 6$. Solving the optimization problem using equation (18), we obtain $\mu = 0.0507$. Substituting this into equation (17), we derive the optimal assignment probabilities shown in Table 23.
Table 23 reflects a solution that balances efficiency and fairness by distributing assignments in a way that respects the cost variance constraint while maximizing the uncertainty in the assignment process.

5. Numerical Illustration

This section presents the applicability of the CUM through real-world data analysis.

Example 7. *We analyze the State/UT-wise data on the number of public and private hospitals enrolled under the Ayushman Bharat–Pradhan Mantri Jan Arogya Yojana (AB-PMJAY) from 2018–19 to 2024–25 using the contingency uniformity measure, in order to compare the distributional patterns between public and private hospitals.*

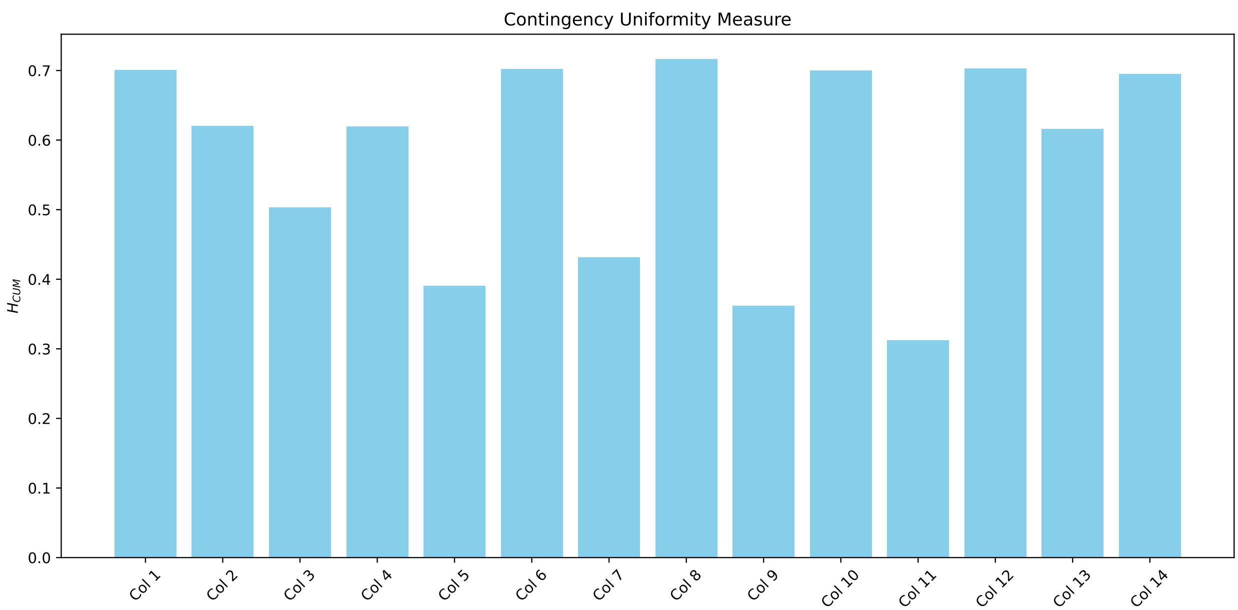

Figure 2. Trends in AB-PMJAY enrollment for public and private hospitals across States/UTs (2018–2025)

In the above plot, odd-numbered columns represent public hospital data, while even-numbered columns correspond to private hospitals. From the visualization, it can be observed that private hospitals exhibit a more uniform pattern of participation in AB-PMJAY across States/UTs compared to public hospitals. This suggests that the distribution of private hospital enrollment is more balanced and less localized.
To validate our claim, a statistical significance test is conducted using two-sample t-test one-tailed at the 5%

Year	**Public** (H_{CUM})	**Private** (H_{CUM})
2018–19	0.7008	0.6202
2019–20	0.5033	0.6192
2020–21	0.3900	0.7021
2021–22	0.4315	0.7163
2022–23	0.3613	0.7001
2023–24	0.3121	0.7024
2024–25	0.6158	0.6949

Table 24. H_{CUM} across years for Public and Private Hospitals

significance level:

$$H_0 : \bar{H}_{CUM(pub)} = \bar{H}_{CUM(priv)} \quad vs \quad H_1 : \bar{H}_{CUM(priv)} > \bar{H}_{CUM(pub)}$$

The mean and variance of the contingency uniformity measures for public($\bar{X}$) and private($\bar{Y}$) hospitals are $\bar{X} = 0.4738$, $\bar{Y} = 0.6793$, $S_X^2 = 0.0236$ and $S_Y^2 = 0.0017$. The test statistics is computed as

$$T_{obs} = \frac{\bar{X} - \bar{Y} - (\bar{H}_{CUM(pub)} - \bar{H}_{CUM(priv)})}{\sqrt{\frac{S_X^2}{n} - \frac{S_Y^2}{m}}}$$

The corresponding p-value is

$$p = P(T > T_{obs}) = 0.0057$$

which is less than the significance level of 0.05. *Hence, we reject the null hypothesis and conclude that the average contingency uniformity measure for private hospitals is significantly higher than that for public hospitals. This implies that private hospitals demonstrate a more uniform and balanced participation pattern across Indian States and Union Territories under the Ayushman Bharat–Pradhan Mantri Jan Arogya Yojana scheme.*

Next, we compare different datasets with varying sizes or category counts.

Example 8. *Suppose we have market-related data for three different products, and the objective is to select one of them for an advertising campaign based on this information.*

Age group ↓/Response →	Buy	Not Buy
Young	10	20
Adult	30	10
Senior	10	20

Table 25. Product 1,$H = 1.6957$, $H_{CUM} = 0.9464$

Age group ↓ /Popularity →	Popular	Not Popular	Neutral
Young	5	15	10
Adult	10	20	10
Senior	15	5	10

Table 26. product 2, $H = 2.1115$, $H_{CUM} = 0.9610$

Although entropy alone cannot be used to directly compare Tables 25, 26, and 27 due to differences in their category structures. The normalized entropy measure H_{CUM} provides a consistent basis for comparison. Among the products, Product 2 exhibits the highest H_{CUM}, indicating a higher degree of uncertainty or variability in its distribution. This higher unpredictability suggests broader and more diverse engagement, making Product 2 a suitable candidate for targeted advertisement efforts.

Age group ↓ /Popularity →	Popular	less Popular	Not Popular	Neutral
Young	5	0	2	3
Adult	10	2	12	1
Senior	3	10	5	9

Table 27. product 3, $H = 2.1738,\ H_{CUM} = 0.8748$

In the above example we clearly see that we have different number of categories for different products, but we can compare them through H_{CUM} for quantification.
A higher CUM represents a distribution that approaches maximum uncertainty a situation where all customer segments or response probabilities are nearly uniform. This can be viewed as a worst-case or benchmark scenario, reflecting a market in which consumers behave randomly and no identifiable segment responds disproportionately better than another. In such a case, a marketer has minimal ability to target specific groups, optimize advertising budgets, or forecast returns. Therefore, maximizing CUM is useful conceptually as a theoretical upper bound that describes the limit of unpredictability. However, in practical marketing applications, the goal may often the opposite. Advertisers usually aim to identify the most predictable, stable, and targetable consumer segments those with lower uncertainty. Thus, one may instead aim to minimize CUM to locate a distribution with the least randomness and the highest concentration of likely responders. A lower CUM indicates that certain customer segments are meaningfully different and therefore offer opportunities for targeted interventions and improved campaign efficiency.
High H_{CUM} models a benchmark scenario of broad unpredictability or wide appeal, which is desirable when selecting a product to promote widely. Low H_{CUM} would be desirable in a different context such as precision targeting or niche segment marketing.

6. Conclusion

We propose a Contingency Uniformity Measure, as a meaningful measure for quantifying data concentration, serving as a robust and normalized entropy-based framework for assessing uncertainty in contingency tables. The fundamental properties of CUM, such as non-negativity, non-expansibility, maximality at the uniform distribution, and weighted additivity, are established to support its theoretical stability. The applicability of CUM is demonstrated through various examples involving different types of two-way tables. Furthermore, we formulate three optimization problems based on CUM under the constraints of fixed marginals, fixed cost, and fixed variability, reflecting realistic scenarios across domains. Appropriate examples are presented to highlight the practical relevance and real-life applications of each formulation.

Acknowledgement

The authors appreciate the comments of the reviewers to improve the quality of the work and the second author gratefully acknowledges the financial support provided by the University Grants Commission (UGC), Government of India, during his Ph.D. work (Student ID: 191620083608).

REFERENCES

1. C. E. Shannon, *A mathematical theory of communication*, The Bell System Technical Journal, vol. 27, no. 3, pp. 379–423, 1948.
2. S. K. Singh, S. Chauhan, A. Alsafrani, M. Islam, H. I. Sherazi, and I. Ullah, *Optimizing healthcare data quality with optimal features driven mutual entropy gain*, Expert Systems, vol. 42, no. 2, p. e13737, 2025.
3. F. Hu, J. Wen, D. Zheng, T. Ying, H. Hou, and W. Wang, *The principle of entropy increase: A novel view of how tourism influences human health*, Journal of Travel Research, vol. 64, no. 3, pp. 752–761, 2025.

4. Y. Sun, J. Liu, K. Yu, M. Alazab, and K. Lin, *PMRSS: Privacy-preserving medical record searching scheme for intelligent diagnosis in IoT healthcare*, IEEE Transactions on Industrial Informatics, vol. 18, no. 3, pp. 1981–1990, 2021.
5. Y. Huang, Y. Zhao, A. Capstick, F. Palermo, H. Haddadi, and P. Barnaghi, *Analyzing entropy features in time-series data for pattern recognition in neurological conditions*, Artificial Intelligence in Medicine, vol. 150, p. 102821, 2024.
6. M. Suguna, R. Kumar, S. Gupta, et al., *Photoplethysmogram (PPG) signal analysis for the heart rate estimation using Shannon entropy*, In: 2025 21st IEEE International Colloquium on Signal Processing & Its Applications (CSPA), pp. 262–265, IEEE, 2025.
7. H. Samawi and A. Helu, *On the inference of entropy measures under different sampling schemes*, Statistics, Optimization & Information Computing, 2025.
8. M. Salajegheh, M. J. Nooghabi and K. Okhli, *A Bayesian method for estimation of the entropy in the presence of outliers based on the contaminated Pareto model*, Statistics, Optimization & Information Computing, vol. 13, no. 5, pp. 1789–1799, 2025.
9. R. Moshwan, X.-L. Shi, W.-D. Liu, J. Liu, and Z.-G. Chen, *Entropy engineering: An innovative strategy for designing high-performance thermoelectric materials and devices*, Nano Today, vol. 58, p. 102475, 2024.
10. J. Liu, Z. Li, D. Lin, Z. Tang, X. Song, P. He, S. Zhang, H. Bian, W. Fu, and Y. Song, *Eutectic high-entropy alloys and their applications in materials processing engineering: A review*, Journal of Materials Science & Technology, vol. 189, pp. 211–246, 2024.
11. M. Bai, W. Qiao, J. Mei, R. Kang, Y. Gao, Y. Wu, Y. Hu, Y. Li, X. Hao, J. Zhao, et al., *High-performance energy storage in $BaTiO_3$-based oxide ceramics achieved by high-entropy engineering*, Journal of Alloys and Compounds, vol. 970, p. 172671, 2024.
12. S. A. S. Fontaine and J. M. Amigó, *Applications of entropy in data analysis and machine learning: A review*, arXiv preprint arXiv:2503.02921, 2025.
13. L. Sun, Z. Huang, H. Peng, Y. Wang, C. Liu, and P. S. Yu, *Lsenet: Lorentz structural entropy neural network for deep graph clustering*, arXiv preprint arXiv:2405.11801, 2024.
14. C. L. Canonne and Q. Yang, *Entropy testing and its application to testing Bayesian networks*, Advances in Neural Information Processing Systems, vol. 37, pp. 132977–133006, 2024.
15. T. M. Cover, *Elements of information theory*, John Wiley & Sons, 1999.
16. N. Kumar, A. Dixit, and V. Vijay, *Entropy measures and their applications: A comprehensive review*, arXiv preprint arXiv:2503.15660, 2025.
17. G. C. Philippatos and C. J. Wilson, *Entropy, market risk, and the selection of efficient portfolios*, Applied Economics, vol. 4, no. 3, pp. 209–220, 1972.
18. E. Maasoumi, *A compendium to information theory in economics and econometrics*, Econometric Reviews, vol. 12, no. 2, pp. 137–181, 1993.
19. R. Ren, Y. Li, Q. Sun, X. Xie, L. Liu, and D. W. Gao, *Digital twin assisted economic dispatch for energy internet with information entropy*, IEEE Transactions on Automation Science and Engineering, vol. 22, pp. 2881–2892, 2025. DOI: 10.1109/TASE.2024.3386358.
20. F. S. Alamri, M. H. Saeed, and M. Saeed, *A hybrid entropy-based economic evaluation of hydrogen generation techniques using multi-criteria decision making*, International Journal of Hydrogen Energy, vol. 49, pp. 711–723, 2024.
21. A. Ochi and Y. Saidi, *Socio-economic inequalities and their impact on children's health and nutrition in Tunisia using generalized entropy measures*, Regional Science Policy & Practice, vol. 16, no. 6, p. 100034, 2024.
22. B. S. Everitt, *The analysis of contingency tables*, CRC Press, 1992.
23. I. J. Good, *Maximum entropy for hypothesis formulation, especially for multidimensional contingency tables*, The Annals of Mathematical Statistics, pp. 911–934, 1963.
24. D. Das and S. Zhou, *Detecting entropy increase in categorical data using maximum entropy distribution approximations*, IISE Transactions, vol. 49, no. 8, pp. 827–837, 2017.
25. D. S. Hooda and M. S. Barak, *Estimation of missing data in design of experiment and contingency table*, SN Applied Sciences, vol. 1, no. 7, p. 670, 2019.
26. H. Gzyl, *Construction of contingency tables by maximum entropy in the mean*, Communications in Statistics—Theory and Methods, vol. 50, no. 20, pp. 4778–4786, 2021.
27. H. Gzyl and S. Mayoral, *Joint probabilities under expected value constraints, transportation problems, maximum entropy in the mean*, Statistica Neerlandica, vol. 78, no. 1, pp. 228–243, 2024.
28. D. Barbará, Y. Li, and J. Couto, *COOLCAT: An entropy-based algorithm for categorical clustering*, In: Proceedings of the Eleventh International Conference on Information and Knowledge Management, pp. 582–589, 2002.
29. C. S. Hong and B. J. Kim, *Mutual information and redundancy for categorical data*, Statistical Papers, vol. 52, no. 1, pp. 17–31, 2011.
30. P. Bogaert and S. Gengler, *MinNorm approximation of MaxEnt/MinDiv problems for probability tables*, AIP Conference Proceedings, vol. 1641, no. 1, pp. 287–296, 2015.
31. E. Skotarczak, A. Dobek, and K. Moliński, *Entropy as a measure of dependency for categorized data*, Biometrical Letters, vol. 55, no. 2, pp. 233–243, 2018.
32. N. Eshima, et al., *Statistical data analysis and entropy*, Springer, 2020.
33. C. Tsallis, *Introduction to nonextensive statistical mechanics: Approaching a complex world*, Springer, 2009.
34. S. Guiasu and A. Shenitzer, *The principle of maximum entropy*, The Mathematical Intelligencer, vol. 7, pp. 42–48, 1985.
35. N. Pesenti, P. Quatto, and E. Ripamonti, *Bootstrap confidence intervals for biodiversity measures based on Gini index and entropy*, Quality & Quantity, vol. 51, no. 2, pp. 847–858, 2017.
36. Ministry of Housing and Urban Affairs, *PC-1/ESCBS Domestic Review Report (09 June 2019)*, Government of India, 2019. Available at: https://mohua.gov.in/upload/uploadfiles/files/PC1_ESCBS_Domestic_Review_Report_%2809_Jun_19%29.pdf.