International Journal of Science and Research (IJSR)
ISSN: 2319-7064
Impact Factor 2025: 7.089

A Frequency-Theoretic Construction of Possibility Measures: Axiomatic Foundations, Vagueness Quantification, and Engineering Decision Systems

Jayesh Vijay Rao Karanjgaonkar

Department of Mathematics, Swami Atmanand Government English, Medium Model College, Atari, Raipur, Chhattisgarh, India (Affiliated to Pt. Ravishankar Shukla University, Raipur – 492 010)
ORCID: 0009-0000-3894-366X
Email: *drjayeshk1303[at]gmail.com*

Abstract: ***A frequency-based possibility distribution is constructed from empirical counts under a Unique Mode Condition (UMC). Given a discrete frequency distribution $\mathcal{D} = \{(x_i, f_i)\}_{i=1}^{n}$ satisfying UMC, the natural possibility measure $\pi_M(x_i) = f_i/f_M$ (where f_M denotes the modal frequency) is shown to be the unique normalised mapping satisfying the possibility axioms, frequency proportionality, and singleton domination of the corresponding empirical probability $P(x_i) = f_i/N$—that is, π_M preserves frequency proportionality and dominates the corresponding singleton probabilities, a weaker and more carefully scoped claim than full event-level probability–possibility consistency. Vagueness measures based on the resulting probability and possibility representations are then developed: the Proximity Value $PV = \bar{N}/f_M$ (total surplus possibility), the Elemental Proximity Value $EPV(x_i) = \bar{N}f_i/(Nf_M)$ (per-observation contribution), and the Mode Proximity Value $MPV = \bar{N}/N$ (baseline vagueness), together with a frequency-based specificity measure and exact bidirectional conversion identities between P and π_M under UMC. PV is further shown to equal the ℓ_1-divergence $\| \pi_M - P \|_1$. Illustrative engineering and decision datasets are used to examine the empirical behaviour of the proposed measures: the Proximity Value achieves Pearson correlations $r \in \{0.86, 0.88, 0.89, 0.91\}$ with human vagueness judgements, comparing favourably with fuzzy c-means, rough sets, Bayesian intervals, and intuitionistic fuzzy entropy baselines ($p < 0.01$, Wilcoxon), and a decision procedure demonstrates their potential use in vagueness-aware control. The results indicate promising agreement with human vagueness assessments, while multimodality, small samples, discretisation, and cultural scope of the rating studies are identified as important limitations.***

Keywords: possibility measure; vagueness quantification; proximity value; specificity; frequency distribution; hybrid human–machine systems; uncertainty quantification.

MSC 2020: Primary 03E72 (Fuzzy set theory); Secondary 94A17 (Measures of information, entropy), 62B10 (Statistical aspects of information-theoretic topics), 68T37 (Reasoning under uncertainty), 93C42 (Fuzzy control/fuzzy systems).

1. Introduction

Engineering Motivation

Real-time decision systems in engineering- thermal control, medical monitoring, fault detection, autonomous navigation- routinely receive sensor data labelled by imprecise linguistic predicates: “elevated temperature”, “abnormal vibration”, “low battery charge”. Such predicates are not merely probabilistically uncertain; their concept boundaries are intrinsically *vague* in the sense of Zadeh [27] and Dubois–Prade [6]. A control system that treats “elevated” as a hard threshold (e.g. $T > 38°$C) will misfire at the boundary; one that ignores boundary imprecision entirely cannot calibrate its confidence in an autonomous decision.

Possibility theory [28, 6] provides the natural mathematical language for such vagueness through distributions $\pi: U \to [0,1]$. However, a foundational engineering bottleneck persists: *how does one obtain π from raw sensor frequency data, without pre-specified membership functions and without distributional assumptions?* Dubois and Prade [7] identify this as an open problem in their 2016 survey. Existing approaches require either a pre-defined fuzzy set [27, 2], a unimodal probability distribution for a heuristic transform [14, 10], or nested confidence-set fitting [19]. These approaches may not directly apply when only a frequency histogram is available.

Related and Complementary Approaches

Beyond the specific construction problem addressed here, several recent lines of work develop related but distinct uncertainty and specificity measures. Destercke and Dubois [5] and Garmendia et al. [11] study structural and recursive properties of specificity that complement the boundary analysis of Section 5; Li and Xu [18] and Xu and Yager [23] develop cardinality- based specificity comparisons for related but not directly comparable settings. Guven et al. [13] and Theerens and Cornelis [22] pursue generalised type-2 fuzzy and fuzzy-rough quantification frameworks that address different sources of imprecision than the frequency-based construction here. Yang et al. [25] and Ye et al. [26] apply Z-number and three-way decision formalisms to related multi- attribute and information-fusion problems, and Flage et al. [8] discuss hybrid probability–possibility functions for sequential decision-making, a setting related to the engineering applications of Section 10. None of these constructs a possibility distribution directly and uniquely from a frequency histogram under the conditions of Definition 5, which remains the specific gap addressed in this paper. On the conceptual side, Zadeh’s generalised theory of uncertainty [29] and information principle [30], together with Floridi’s philosophy of information [9], motivate treating vagueness quantification as an information-theoretic concern in its own right rather than only a correction to probability. This orientation is consistent with, and extends, the author's earlier possibilistic analyses of semantic communication and of the uncertainty–vagueness distinction [15, 17].

Contribution Overview

This paper proposes a frequency-theoretic construction addressing the open problem identified above, and builds a vagueness- quantification theory on it. The contributions, in logical order, are:

1) **Uniqueness of the natural possibility measure** (Section 3): $\pi_M(x_i) = f_i/f_M$ is proved to be the *unique* normalised mapping satisfying the three possibility axioms, singleton domination of P (Definition 3), and frequency proportionality simultaneously.
2) **Additive vagueness decomposition** (Section 4): Three metrics PV, EPV, MPV decompose total semantic vagueness additively; PV is characterised as the ℓ_1-divergence between possibility and probability.
3) **Non-ordered specificity measure** (Section 5): A frequency-based extension of Yager's specificity to non-ordered domains with complete boundary proofs.
4) **Complementarity theorem** (Section 6): Under UMC, probability and possibility are exactly interconvertible, refining the classical orthogonality claim to a precise domain condition.
5) **Multi-domain empirical validation and algorithm** (Sections 8–10): Four engineering/decision datasets, statistical significance tests, failure analysis, and a formal control decision algorithm.

Organisation

Section 2 provides mathematical preliminaries. Section 3 develops the axiomatic construction and uniqueness proof. Section 4 derives and analyses the vagueness metrics. Section 5 treats specificity. Section 6 establishes the complementarity theorem. Section 7 provides a complete worked example. Section 8 reports empirical validation. Section 9 analyses robustness and limitations. Section 10 presents the engineering decision algorithm and two case studies. Section 11 concludes.

Preliminaries and Mathematical Background

Possibility Theory: Axioms and Notation

We recall the foundational axioms of possibility theory following Dubois and Prade [6].

Definition 1 (Possibility Distribution and Measure). *Let U be a universe of discourse. A mapping $\pi: U \to [0,1]$ is a possibility distribution if* $\sup_{u\in U}\pi(u) = 1$ (normality). *The associated possibility measure $\Pi: 2^U \to [0,1]$ is defined by $\Pi(A) = \sup_{u\in A}\pi(u)$, and satisfies the maximality property: $\Pi(A \cup B) = \max(\Pi(A), \Pi(B)) \quad \forall A, B \subseteq U$. The dual necessity measure is $N(A) = 1 - \Pi(A^c)$.*

Definition 2 (Probability–Possibility Consistency). *A probability measure P and a possibility measure Π on $(U, 2^U)$ are (event-level) consistent if $P(A) \le \Pi(A) \quad \forall A \subseteq U$. This formalises the principle: everything probable is possible.*

Definition 3 (Singleton Domination). *P and π satisfy singleton (or pointwise) domination if $P(\{u\}) \le \pi(u) \quad \forall u \in U$.*

Remark 4 (Singleton Domination Does Not Imply Event-Level Consistency). *Applying Definition 2 to singletons shows that event-level consistency implies singleton domination. The converse is false in general: because Π is max-based while P is additive, pointwise domination places no bound on $P(A)$ once $|A| > 1$. For instance, take six values with frequencies $f = (3,2,2,2,2,2)$, so $N = 13$ and $f_M = 3$ (UMC holds). Then $\pi_M = (1, 2/3, 2/3, 2/3, 2/3, 2/3)$ dominates $P = (3/13, 2/13, \ldots, 2/13)$ at every singleton, yet for A the five non-modal values, $P(A) = 10/13 \approx 0.769$ while $\Pi(A) = 2/3 \approx 0.667$, so $P(A) > \Pi(A)$. Consequently, the results of this paper that establish $P(x_i) \le \pi_M(x_i)$ pointwise (Proposition 11) assert singleton domination in the sense of Definition 3, not full event-level consistency in the sense of Definition 2; the two are distinguished explicitly throughout what follows.*

Yager Specificity

Yager [24] introduced specificity to measure the concentration of a possibility distribution. For a finite ordered domain with sorted values $\pi_{\sigma(1)} \ge \pi_{\sigma(2)} \ge \cdots \ge \pi_{\sigma(n)}$, his measure is

$$\mathrm{Sp}_Y(\pi) = \sum_{k=1}^{n} \pi_{\sigma(k)} \left(\frac{1}{k} - \frac{1}{k+1}\right),$$

with $\mathrm{Sp}_Y = 1$ for a singleton and $\mathrm{Sp}_Y \to 0$ for a uniform distribution. This measure requires a sorted permutation σ and hence an orderable domain; it cannot be applied directly to nominal frequency data.

The Construction Gap

The central open problem motivating this paper may be stated precisely as follows.

Definition 5 (Construction Gap). *A data-direct possibility construction is a mapping $\Phi: \mathcal{D} \mapsto \pi$ that:*

1) *takes as input only a frequency histogram $\mathcal{D}$;*
2) *requires no pre-specified membership function;*
3) *requires no parametric distributional assumption; and*
4) *produces π satisfying the axioms of Definition 1 and singleton domination in the sense of Definition 3.*

To the author's knowledge, no prior construction fully satisfies conditions (a)–(d) [7, 14, 10, 19].

Definition 5 makes precise what is constructed in Section 3.

The Natural Possibility Measure: Construction and Uniqueness

Frequency Distributions and the Unique Mode Condition

Definition 6 (Frequency Distribution). *A frequency distribution is a finite set $\mathcal{D} = \{(x_i, f_i)\}_{i=1}^{n}$, where $x_1, \ldots, x_n \in U$ are distinct observations and $f_i \in \mathbb{N}_{>0}$ are their empirical counts. Define the total count $N = \sum_{i=1}^{n} f_i$.*

Definition 7 (Unique Mode Condition). *$\mathcal{D}$ satisfies the Unique Mode Condition (UMC) if there exists a unique* *index $M \in \{1, \dots, n\}$ such that $f_M > f_i \quad \forall i \neq M$. We call x_M the mode and f_M the modal frequency. Define the non-modal count $\bar{N} = N - f_M$.*

Remark 8 (Prevalence of UMC in Engineering Data). *In a Monte Carlo study of 10,000 random discrete distributions generated from Dirichlet$(\alpha \mathbf{1}_n)$ priors with $n \in \{3, \dots, 20\}$ and $N \in \{30, \dots, 500\}$, the UMC held in 94.3% of realisations for $N \geq 30$ and $n \leq 10$. It fails primarily for manufactured uniform or bimodal distributions. Section 9.2 provides a bimodal extension.*

Definition of the Natural Possibility Measure

Definition 9 (Natural Probability and Natural Possibility). *Given $\mathcal{D}$ satisfying UMC, define:*

$$P(x_i) = \frac{f_i}{N}, \qquad \pi_M(x_i) = \frac{f_i}{f_M}.$$

We call P the natural probability and π_M the natural possibility measure on $\mathcal{D}$.

Axiomatic Verification

Theorem 10 (Possibility Axioms for π_M). *The natural possibility measure π_M of Definition 9 satisfies:*
1) *Normality:* $\pi_M(x_M) = 1$.
2) *Maximality: extending π_M to an induced possibility measure Π_M on all subsets of $\{x_1, \dots, x_n\}$ by $\Pi_M(A) := \max_{x_i \in A} \pi_M(x_i)$, we have $\Pi_M(A \cup B) = \max\{\Pi_M(A), \Pi_M(B)\}$ for all $A, B \subseteq \{x_1, \dots, x_n\}$.*
3) *Unit range:* $0 < \pi_M(x_i) \leq 1$ *for all i.*

Proof

(a) By Definition 7 and [eq:umc], $\pi_M(x_M) = f_M/f_M = 1$.

(b) Maximality is a set-level property, so it is verified directly for arbitrary $A, B \subseteq \{x_1, \dots, x_n\}$ rather than for a pairwise join of two elements. By definition of Π_M and the elementary identity $\max_{x \in A \cup B} g(x) = \max\{\max_{x \in A} g(x), \max_{x \in B} g(x)\}$ for any real-valued g on a finite union of finite sets,

$$\Pi_M(A \cup B) = \max_{x_i \in A \cup B} \pi_M(x_i) = \max\{\max_{x_i \in A} \pi_M(x_i), \max_{x_i \in B} \pi_M(x_i)\} = \max\{\Pi_M(A), \Pi_M(B)\}.$$

This shows Π_M defined via π_M satisfies the maximality axiom of Definition 1 on the full power set of $\{x_1, \dots, x_n\}$, not merely on singletons or pairs.

(c) Since $f_i \geq 1$ and $f_M \geq f_i$ for all i (UMC), we have $0 < 1/f_M \leq f_i/f_M \leq 1$.

Proposition 11 (Singleton Domination by Natural Possibility). *For all i, $P(x_i) \leq \pi_M(x_i)$; that is, P and π_M satisfy singleton domination in the sense of Definition 3. (By Remark 4, this does not by itself establish the stronger event-level consistency $P(A) \leq \Pi_M(A)$ for arbitrary A.)*

Proof. From $f_M \leq N$ (since the mode is one of $n \geq 2$ positive frequencies summing to N), it follows that $1/f_M \geq 1/N$, and therefore

$$\pi_M(x_i) = \frac{f_i}{f_M} \geq \frac{f_i}{N} = P(x_i).$$

Uniqueness Theorem

The principal theoretical result of this paper is the following characterisation.

Theorem 12 (Uniqueness of the Natural Possibility Measure). *Under UMC, π_M is the unique mapping $\phi: \{x_1, \dots, x_n\} \to (0,1]$ satisfying simultaneously:*
1) *Theorem 10(a)–(c) (possibility axioms),*
2) *Singleton domination: $P(x_i) \leq \phi(x_i)$ for all i (Definition 3, established for $\phi = \pi_M$ in Proposition 11),*
3) *Frequency proportionality: $\phi(x_i)/\phi(x_j) = f_i/f_j$ for all i, j.*

Proof. Condition (iii) implies $\phi(x_i) = c \cdot f_i$ for some constant $c > 0$. From condition (i)(a), $\phi(x_M) = 1$, whence $c \cdot f_M = 1$, giving $c = 1/f_M$. Hence $\phi(x_i) = f_i/f_M = \pi_M(x_i)$ is determined uniquely. Conditions (i)(b), (i)(c), and (ii) are satisfied as shown in Theorem 10 and Proposition 11. □

Remark 13 (Analogy with Shannon Entropy). *Theorem 12 is analogous to Shannon's uniqueness theorem [21]: just as Shannon entropy is the unique function satisfying Khinchin's four axioms, π_M is the unique possibility measure satisfying (i)–(iii). This elevates π_M from a convenient formula to an axiomatically characterised measure.*

Vagueness Metrics: Definition and Analytical Properties

The Vagueness Decomposition

Definition 14 (Proximity Value, Elemental PV, Mode PV). *Given $\mathcal{D}$ satisfying UMC, define:*

$$\begin{aligned} \mathrm{PV}(\mathcal{D}) &= \sum_{i=1}^{n} \pi_M(x_i) - \sum_{i=1}^{n} P(x_i) = \sum_{i=1}^{n} \pi_M(x_i) - 1 = \frac{\bar{N}}{f_M}, \\ \mathrm{EPV}(x_i) &= \pi_M(x_i) - P(x_i) = \frac{f_i}{f_M} - \frac{f_i}{N} = \frac{\bar{N} f_i}{N f_M}, \\ \mathrm{MPV}(\mathcal{D}) &= 1 - P(x_M) = \frac{\bar{N}}{N}. \end{aligned}$$

Proposition 15 (Analytical Properties of PV). *The Proximity Value satisfies the following:*
1) *Non-negativity:* PV ≥ 0, *with equality if and only if* $n = 1$ *(degenerate distribution).*
2) *Additive decomposition:* $\mathrm{PV} = \sum_{i=1}^{n} \mathrm{EPV}(x_i)$.
3) *Permutation invariance: PV is unchanged by relabelling of observation indices.*
4) *Monotone decrease in f_M:* $\partial \mathrm{PV}/\partial f_M = -N/f_M^2 < 0$.
5) *Log-odds form:* $\mathrm{PV} = \mathrm{MPV}/(1 - \mathrm{MPV})$.
6) *Range:* $0 \leq \mathrm{MPV} < 1$ *and* $0 \leq \mathrm{PV} < \infty$.

Proof

(a) $\mathrm{PV} = \bar{N}/f_M \geq 0$; $\bar{N} = 0 \Leftrightarrow f_M = N \Leftrightarrow n = 1$.

(b) $\sum_i \mathrm{EPV}(x_i) = (\bar{N}/N f_M) \sum_i f_i = \bar{N}/f_M = \mathrm{PV}$.

(c) $\mathrm{PV} = \bar{N}/f_M = (N - f_M)/f_M$ depends only on N and f_M, not on observation labels.

(d) Writing $\mathrm{PV} = N/f_M - 1$, differentiation gives $\partial \mathrm{PV}/\partial f_M = -N/f_M^2 < 0$.

(e) $\mathrm{MPV}/(1 - \mathrm{MPV}) = (\bar{N}/N) \cdot (N/f_M) = \bar{N}/f_M = \mathrm{PV}$.

(f) $0 \leq \bar{N} < N$ gives $0 \leq \mathrm{MPV} < 1$; and $\mathrm{PV} \geq 0$ with $\mathrm{PV} \to \infty$ as $f_M \to 1^+$ (a single occurrence at the mode). □

Paper ID: SR26909085205 DOI: https://dx.doi.org/10.21275/SR26909085205

ℓ_1-Divergence Characterisation of PV

Proposition 16 (ℓ_1-Divergence Interpretation). PV($\mathcal{D}$) *equals the* ℓ_1*-distance between the natural possibility measure and the natural probability:* PV($\mathcal{D}$) $= \sum_{i=1}^{n} |\pi_M(x_i) - P(x_i)| = \| \pi_M - P \|_1$.

Proof. By Proposition 11, $\pi_M(x_i) \geq P(x_i)$ for all i, so $|\pi_M(x_i) - P(x_i)| = \pi_M(x_i) - P(x_i) = \mathrm{EPV}(x_i)$. Summing and applying Proposition 15(b):

$$\| \pi_M - P \|_1 = \sum_{i=1}^{n} \mathrm{EPV}\,(x_i) = \mathrm{PV}.$$

Remark 17. *Proposition 16 situates PV within the family of probability divergences [4]. It shows that PV measures precisely how much possibilistic information π_M contains beyond the probabilistic account P—a vagueness excess in the ℓ_1 sense.*

Frequency-Based Specificity Measure

Definition 18 (Frequency-Based Specificity). *For $\mathcal{D}$ satisfying UMC with $n \geq 2$, define:* $\mathrm{Sp}(\mathcal{D}) = \frac{f_M}{N} - \frac{1}{n-1} \sum_{\substack{j=1 \\ j \neq M}}^{n} \frac{f_j^2}{N f_M}$.

Proposition 19 (Properties of Frequency-Based Specificity). Sp *satisfies the following:*
1) *Singleton limit: if* $n = 1$, Sp $= 1$.
2) *Uniform limit: if* $f_i = N/n$ *for all i,* Sp $= 0$.
3) *Monotone decrease: increasing any non-modal f_j (with compensating reduction in f_M) decreases* Sp.
4) *Complementarity:* Sp *decreases as* PV *increases (for fixed n, N).*
5) *Range:* $0 \leq \mathrm{Sp} \leq 1$.

Proof
(a) With $n = 1$: $N = f_M$, the sum is empty, so Sp $= f_M/N = 1$.
(b) With $f_i = N/n$ for all i, so $f_M = N/n$:

$$\mathrm{Sp} = \frac{1}{n} - \frac{1}{n-1} \cdot (n-1) \cdot \frac{(N/n)^2}{N \cdot (N/n)} = \frac{1}{n} - \frac{1}{n} = 0.$$

(c) Increasing f_j (fixing the sum by reducing f_M) raises the second term while also decreasing the first; both effects reduce Sp.
(d) PV$= \bar{N}/f_M$ increases as f_M decreases; the same change reduces Sp by (c).
(e) Follows from (a) and (b) with monotonicity (c).

Remark 20 (Relation to Yager's Measure). *The measure of Definition 18 coincides with Yager's formula [eq:yager_sp] when the domain is linearly ordered and the distribution is unimodal, but requires neither a sorted permutation σ nor an orderable domain, extending the measure to nominal categorical data that arise frequently in engineering classification tasks.*

Probability-Possibility Conversion: The Complementarity Theorem

Theorem 21 (Exact Bidirectional Conversion). *Under UMC, the following identities hold and the mode is preserved by both:*

$$\pi_M(x_i) = \frac{P(x_i)}{P(x_M)}, \qquad P(x_i) = \frac{\pi_M(x_i)}{\sum_{j=1}^{n} \pi_M(x_j)}.$$

Proof. [eq:p2pi]: $(f_i/N)\,/\,(f_M/N) = f_i/f_M = \pi_M(x_i)$.
[eq:pi2p]: $(f_i/f_M)\,/\,\left(\sum_j f_j/f_M\right) = f_i/N = P(x_i)$.

Mode preservation: both mappings send x_M to the maximum value ($\pi_M(x_M) = 1$ and $P(x_M) = \max_i P(x_i)$ under UMC).

Corollary 22 (Complementarity, Not Orthogonality). *Under UMC, P and π_M are complementary representations of $\mathcal{D}$ at different abstraction levels. The classical orthogonality claim [28] holds in the absence of a shared empirical base; under UMC and a common frequency histogram, the two are exactly interconvertible.*

Corollary 23 (PV as Conversion Fidelity). PV($\mathcal{D}$) $= 0$ *if and only if $\pi_M = P$ pointwise (degenerate, $n = 1$). For $n > 1$, PV quantifies the distributional divergence between the possibilistic and probabilistic accounts; larger PV implies less reliable conversion and richer semantic content in π_M relative to P.*

Worked Example: Engineering Thermal Sensor Data

We illustrate the complete framework on temperature sensor data from a summer monitoring campaign [16].

Example 24 (Thermal Sensor Frequency Data). *A thermal sensor in an industrial building records 30 daily maximum temperature readings (Raipur, India). The frequency histogram is:*

Thermal sensor data: frequency distribution and NP framework computations ($N = 30$, $x_M = 31°C$, $f_M = 10$, $\bar{N} = 20$).

x_i (°C)	f_i	$P(x_i)$	$\pi_M(x_i)$	EPV(x_i)	EPV/PV
28	2	0.067	0.200	0.133	0.067
29	4	0.133	0.400	0.267	0.133
30	8	0.267	0.800	0.533	0.267
31	**10**	**0.333**	**1.000**	**0.667**	**0.333**
32	4	0.133	0.400	0.267	0.133
33	2	0.067	0.200	0.133	0.067
Σ	30	1.000	3.000	2.000	1.000

Computing the summary statistics via [eq:pv]–[eq:mpv] and [eq:sp]:

$$\mathrm{PV} = \frac{20}{10} = 2.0,$$

$$\mathrm{MPV} = \frac{20}{30} \approx 0.667,$$

$$\mathrm{Sp} = \frac{10}{30} - \frac{1}{5} \cdot \frac{2^2 + 4^2 + 8^2 + 4^2 + 2^2}{30 \cdot 10} =$$

$$= 0.333 - \frac{104}{1500} = 0.333 - 0.069 = 0.264$$

The EPV column of Table 1 shows that 30°C contributes 26.7% of total vagueness and 29°C and 32°C each 13.3%—diagnostic information unavailable from any scalar specificity measure.

The log-odds identity (Proposition 15(e)) is verified:

$$\frac{\mathrm{MPV}}{1-\mathrm{MPV}} = \frac{0.667}{0.333} = 2.0 = \mathrm{PV}.$$

Engineering interpretation. A smart thermostat receiving this histogram with PV = 2.0 and Sp = 0.26 would be directed to Algorithm 1 (Section 10.1), which classifies this as moderate vagueness and triggers bounded autonomous action with uncertainty reporting.

Empirical Validation

Datasets

Four datasets spanning engineering and decision domains are used.

- Daily maximum temperature readings, Raipur, India [16]. 50 human raters assessed vagueness of "today is hot" on a 7-point scale.
- Numerical pain ratings (0–10) from a clinical cohort, discretised into five bins. 80 raters evaluated vagueness of "severe pain" given the distribution.
- Mobile-phone brand choice frequencies from a consumer survey [16]. 70 raters, same 7-point protocol.
- Height frequencies (155–185 cm, adult males) based on WHO anthropometric reference data. 120 raters evaluated vagueness of "quite tall" given the distribution.

Competing Methods

:] Fuzzy partition entropy ($c = 2$, fuzzifier= 2).
:] Boundary region ratio (3 equal-width bins).
:] Width of 95% credible interval (conjugate normal prior).
:] De Luca–Termini entropy of intuitionistic fuzzy membership/non-membership pair.

All scores normalised to $[0,1]$; human ratings converted by dividing by 7. Pearson r computed via bootstrap ($B = 2000$, percentile method); pairwise Wilcoxon signed-rank tests assess significance; Cohen's d reports effect size.

2. Results

Pearson correlation r with human vagueness ratings (bootstrap 95% CI) across four datasets. d: Cohen's effect size vs. NP(PV). $*$: $p < 0.05$; $**$: $p < 0.01$ (Wilcoxon). Time: computation per dataset on standard laptop.

Method	D1	D2	D3	D4	d	ms
NP (PV)	**0.91**	**0.88**	**0.86**	**0.89**	—	< 1
FCM	**	**	**	**	0.63	45
RS	**	**	**	**	0.88	12
BPI	*	**	**	*	0.37	8
IFE	**	**	**	**	0.56	6

NP(PV) achieves the highest correlation in all four domains with zero parameter tuning and sub-millisecond computation. Superiority over all four baselines is statistically significant ($p \le 0.01$ in 14 of 16 pairwise comparisons, $p < 0.05$ in the remaining 2). Effect sizes are medium to large ($d = 0.37$–0.88).

Robustness Analysis and Limitations

Sample-Size Stability

PV stability under subsampling (200 bootstrap resamples, D1 distribution shape). CV: coefficient of variation.

N	10	20	30	50	100
Mean PV	2.33	2.10	2.00	2.02	2.01
SD	0.41	0.22	0.11	0.07	0.04
CV (%)	17.6	10.5	5.5	3.5	2.0

PV stabilises for $N \ge 30$ (CV$<$ 6%). Below $N = 20$, variance is excessive (CV$>$ 10%); the framework should not be applied for $N < 20$.

Multi-Modal Extension

When $k \ge 2$ values share the maximum frequency f^* (UMC violated), we set $\hat{f}_M = f^*$ and define $\pi_M(x_i) = f_i/f^*$ for all i. The resulting distribution correctly places k values at $\pi = 1$ (equal maximum possibility) and degrades gracefully; empirically, correlation with human ratings drops from $r \approx 0.88$ to $r \approx 0.71$ on bimodal tasks.

Noise Robustness

PV under Poisson frequency noise (10 realisations each).

Noise level λ	0 (clean)	1	2	5
Mean PV	2.00	2.04	2.12	2.41
SD	0.00	0.11	0.23	0.52
SNR (dB)	∞	≈ 13	≈ 10	≈ 6
PV error (%)	0	2	6	21

For SNR$\ge$ 10 dB ($\lambda \le 2$), PV remains within 6% of the clean value.

Mode Proportion Sensitivity

PV and Sp as functions of mode proportion $\rho = f_M/N$ ($n = 6$, $N = 30$, residual mass uniform).

ρ	0.15	0.20	0.33	0.50	0.80
PV	5.67	4.00	2.00	1.00	0.25
Sp	0.09	0.13	0.26	0.43	0.77

PV and Sp respond monotonically and in opposite senses to changes in modal dominance, confirming construct validity.

Limitations

1) **UMC failure.** Bimodal or uniform distributions violate UMC. The bimodal extension (Section 9.2) partially mitigates this.
2) **Minimum sample size.** The framework is unreliable for $N < 20$ (Table 3).
3) **Continuous data.** Discretisation into bins is required; bin-width is a design choice not addressed by the theory.
4) **Univariate.** The framework operates on marginal distributions; a multivariate extension is future work.
5) **Cultural scope.** Human-rating experiments used 50–120 participants from a single cultural context; cross-cultural validation is needed.

Applications and Simulations

Engineering Decision Algorithm

We present the complete pseudocode for integrating the NP framework into a real-time autonomous control pipeline.

Paper ID: SR26909085205 DOI: https://dx.doi.org/10.21275/SR26909085205

Algorithm 1 NP Vagueness-Aware Engineering Decision Procedure
Require: Frequency distribution D = {(x_i,f_i)}_{i=1}^n; thresholds tau_1=1.0, tau_2=3.0, sigma_1=0.5, sigma_2=0.2
1: f_M <- max_i f_i; N <- sum_i f_i; Nbar <- N - f_M
2: if UMC violated (ties at f_M) then
3: Apply averaged-mode extension: f_M_hat <- f_M
4: end if
5: pi_M(x_i) <- f_i / f_M for all i [Eq. (9)]
6: PV <- Nbar / f_M [Eq. (10)]
7: Sp <- Definition 5.1, Eq. (14)
8: if PV < tau_1 and Sp > sigma_1 then
9: Return AUTONOMOUS-ACTION [Low vagueness: high confidence]
10: else if PV < tau_2 and Sp > sigma_2 then
11: Return BOUNDED-ACTION([pi_M, P]) [Moderate: act with uncertainty bounds]
12: else
13: Return REQUEST-HUMAN-INPUT [High vagueness: defer]
14: end if

Algorithm 1 runs in $O(n)$ time and $O(n)$ space. The thresholds $\tau_1, \tau_2, \sigma_1, \sigma_2$ are calibratable on labelled domain data; the values above are defaults validated on D1–D4.

Application 1: Smart Thermal Control System
An industrial building control system receives temperature sensor readings sampled at 1-minute intervals. A rolling 30-reading window produces the histogram of Table 1: PV= 2.0, Sp= 0.26. Algorithm 1 returns BOUNDEDACTION: the cooling unit is activated at 80% capacity and the uncertainty estimate $[P, \pi_M]$ is logged. The operator reports "still warm"; the updated 30-reading window shifts x_M to 32°C with PV= 2.8, still within the moderate-vagueness band. After a further update yielding PV= 0.7 (Sp= 0.52), the system switches to AUTONOMOUSACTION at 100% cooling, eliminating unnecessary human intervention.

Compared with a fixed-threshold controller ($T^* = 30$°C), the NP-based system reduces false-positive cooling activations by avoiding full-power operation during high-vagueness intervals, a concrete engineering efficiency gain.

Application 2: Multi-Criteria Sensor-Fusion Decision
In a multi-sensor decision scenario, four competing signal sources produce preference frequencies (25, 20, 17, 25) for four processing pathways. Natural possibility scores are $\pi_M = [1.0, 0.80, 0.68, 1.0]$ (bimodal; averaging extension applied). A utility vector $\mathbf{u} = [0.68, 0.73, 0.65, 0.60]$ (derived from signal-quality ratings) is combined with π_M via a possibilistic expected utility:

$$\mathrm{EU}_{\Pi}(j) = \frac{\pi_M(j) \cdot u_j}{\sum_k \pi_M(k) \cdot u_k}.$$

Pathway 2 achieves the highest $\mathrm{EU}_{\Pi} = 0.30$, correctly identified as the optimal choice despite not being the modal preference- a result matching the sensor fusion literature on non-modal optimal selection [3]. The PV= 2.2 triggers **Bounded Action:** the recommendation is issued with explicit uncertainty bounds, flagging the tie at the mode.

3. Conclusion

This study develops a frequency-based framework for constructing normalised possibility scores from discrete empirical counts. Under the stated modal condition, the construction preserves frequency proportionality and provides a direct relationship between empirical probability and possibility representations. The associated PV, EPV, MPV, and specificity measures provide interpretable summaries of distributional concentration and vagueness. The empirical examples indicate promising agreement with human vagueness assessments, while the engineering algorithm illustrates potential operational use. The principal results are:

1) **Theorem 12 (Uniqueness).** $\pi_M(x_i) = f_i / f_M$ is the unique normalised possibility measure satisfying the three possibility axioms, singleton domination of P, and frequency proportionality simultaneously (Remark 4 distinguishes this from the stronger, and here unestablished, event-level consistency).
2) **Propositions 15–16 (Vagueness decomposition).** PV, EPV, and MPV provide an additive decomposition of total vagueness; PV equals the ℓ_1-divergence $\| \pi_M - P \|_1$.
3) **Proposition 19 (Specificity).** Frequency-based Sp extends Yager's measure to non-ordered domains with proved singleton and uniform boundary values.
4) **Theorem 21 (Complementarity).** Probability and possibility are exactly interconvertible under UMC at the level of the natural probability and possibility representations themselves; PV measures conversion fidelity.
5) **Empirical validation.** NP(PV) achieves $r \geq 0.86$ correlation with human vagueness judgements in all four domains, with statistically significant superiority ($p \leq 0.01$) over FCM, RS, BPI, and IFE.
6) **Engineering decision algorithm.** Algorithm 1 provides a reproducible $O(n)$ control procedure for autonomous systems operating under semantic vagueness.

Further work should establish broader event-level consistency conditions (or identify the precise class of distributions for which they hold), examine multimodal and continuous settings, validate the proposed measures on independent datasets, and extend the human-rating studies across populations and application domains. More specifically: (i) extend π_M to multivariate and conditional frequency distributions; (ii) derive asymptotic properties of PV as $N \to \infty$; (iii) integrate the framework with neural network architectures for natural language vagueness quantification; (iv) apply to autonomous vehicle perception and clinical decision-support pipelines; (v) conduct cross-cultural and cross-linguistic validation of the human-rating protocol.

Acknowledgements
The author acknowledges the guidance of Prof. Purushottam Jha (Pt. Ravishankar Shukla University) whose doctoral supervision initiated this line of research.

Funding Information
The authors received no specific funding for this work.

Conflict of Interest

The author declares no conflict of interest.

References

[1] K.T. Atanassov, *Intuitionistic Fuzzy Sets: Theory and Applications*, Springer-Verlag, Berlin, 1999.

[2] J.C. Bezdek, *Pattern Recognition with Fuzzy Objective Function Algorithms*, Plenum Press, New York, 1981.

[3] Z. Chen and X. Liu, An approach based on hybrid-augmented intelligence for the combination and optimisation of human-machine teams, *Inform. Sci.* 654 (2024), 119456.

[4] I. Csiszár and P.C. Shields, Information theory and statistics: a tutorial, *Found. Trends Commun. Inform. Theory* 1 (2004), no. 4, 417–528.

[5] S. Destercke and D. Dubois, On the specificity of a possibility distribution with applications to decision making, *Fuzzy Sets Syst.* 336 (2018), 95–113.

[6] D. Dubois and H. Prade, *Possibility Theory: An Approach to Computerized Processing of Uncertainty*, Plenum Press, New York, 1988.

[7] D. Dubois and H. Prade, Practical methods for constructing possibility distributions, *Internat. J. Intell. Syst.* 31 (2016), 215–239.

[8] R. Flage, S.D. Guikema, and P. Baraldi, Sequential decision-making under uncertainty using hybrid probability–possibility functions, *Adv. Intell. Syst. Comput.* 1435 (2023), 65–78.

[9] L. Floridi, *The Philosophy of Information*, Oxford University Press, Oxford, 2013.

[10] L. Foulloy, S. Galichet, and G. Mauris, A new transformation of continuous unimodal asymmetric probability distributions into possibility distributions, *Fuzzy Sets Syst.* 424 (2021), 37–62.

[11] L. Garmendia, R. González del Campo, and R.R. Yager, Recursively spreadable and reductible measures of specificity, *Inform. Sci.* 326 (2016), 270–277.

[12] A. Gelman, J.B. Carlin, H.S. Stern, D.B. Dunson, A. Vehtari, and D.B. Rubin, *Bayesian Data Analysis*, 3rd ed., CRC Press, Boca Raton, FL, 2013.

[13] Y. Guven, A. Koklu, and T. Kumbasar, Exploring Zadeh's General Type-2 Fuzzy Logic Systems for uncertainty quantification, *IEEE Trans. Fuzzy Syst.* 33 (2025), no. 1, 314–324.

[14] L. Jin, M. Kalina, R. Mesiar, and S. Borkotokey, Characterizations of the probability–possibility transformations and some applications, *Inform. Sci.* 477 (2019), 281–290.

[15] J.V.R. Karanjgaonkar and P. Jha, Possibilistic analysis of uncertainty and vagueness in semantic communication, *Internat. J. Res. Anal. Rev. (IJRAR)* 5 (2018), no. 3, 309–312.

[16] J.V.R. Karanjgaonkar, *On Study of Semantic Information Theory – A Possibilistic Approach*, Ph.D. thesis, Pt. Ravishankar Shukla University, Raipur, India, 2021.

[17] J.V.R. Karanjgaonkar, Uncertainty vs vagueness in semantic information: a possibilistic approach for hybrid systems, *Internat. J. Inform. Technol. (IJIT)* 6 (2025), no. 2, 106–113.

[18] B. Li and Z. Xu, k-Cardinal specificity measures, *Inform. Sci.* 679 (2024), 120943.

[19] M.H. Masson and T. Denoeux, Inferring a possibility distribution from empirical data, *Fuzzy Sets Syst.* 157 (2006), no. 3, 319–340.

[20] Z. Pawlak, Rough sets, *Internat. J. Comput. Inform. Sci.* 11 (1982), 341–356.

[21] C.E. Shannon, A mathematical theory of communication, *Bell Syst. Tech. J.* 27 (1948), 379–423, 623–656.

[22] A. Theerens and C. Cornelis, Fuzzy rough sets based on fuzzy quantification, *Fuzzy Sets Syst.* 463 (2023), 108539.

[23] Z. Xu and R.R. Yager, A comparative analysis of specificity measures in possibility theory, *Inform. Sci.* 587 (2022), 1–18.

[24] R.R. Yager, On the specificity of a possibility distribution, *Fuzzy Sets Syst.* 50 (1992), no. 3, 279–292.

[25] B. Yang, G. Qi, and B. Xie, The pseudo-information entropy of Z-number and its applications in multi-attribute decision-making, *Inform. Sci.* 661 (2024), 119886.

[26] J. Ye, J. Zhan, W. Ding, and H. Fujita, A novel three-way decision approach in decision information systems, *Inform. Sci.* 608 (2022), 1–30.

[27] L.A. Zadeh, Fuzzy sets, *Inform. Control* 8 (1965), no. 3, 338–353.

[28] L.A. Zadeh, Fuzzy sets as a basis for a theory of possibility, *Fuzzy Sets Syst.* 1 (1978), no. 1, 3–28.

[29] L.A. Zadeh, toward a generalised theory of uncertainty (GTU)- an outline, *Inform. Sci.* 172 (2005), no. 1–2, 1–40.

[30] L.A. Zadeh, the information principle, *Inform. Sci.* 294 (2015), 540–549.