Pure Mathematics For Theoretical Computer Science (PMTCS) *Vol. 06, No. 01, PP. 01-21, 2026*

An Introduction to Probability, Hyper-Probability, and Super-Hyper-Probability

Takaaki Fujita[1,*], **Ajoy Kanti Das**[2]
[1]Independent Researcher, Tokyo, Japan
[2]Associate Professor, Department of Mathematics, Tripura University, Agartala-799022, Tripura, India
Emails: Takaaki.fujita060@gmail.com; ajoykantidas@gmail.com

Abstract

Standard probability theory assigns each event a single real value in $[0, 1]$, satisfying non-negativity, normalization, and countable additivity. Hyper-Probability extends this notion by assigning to each event a set of probability values in $[0, 1]$, thereby capturing multiple independent assessments from diverse sources. Super-HyperProbability further generalizes the framework by mapping events to iterated power sets of $[0, 1]$, modeling hierarchical uncertainty across multiple aggregation levels. In this paper, we formally define the Hyper-Probability Measure and Hyper-Probability Distribution, examine their fundamental properties, and demonstrate how these constructs unify and extend classical probability within the Hyper- and Super-HyperProbability paradigms.

Keywords: Probability; HyperProbability; SuperHyperProbability; Probability Distribution; Probability Measure

1 Preliminaries

This section provides an introduction to the foundational concepts and definitions required for the discussions in this paper.

1.1 Power Set and nth Power Set

Power set of set S is the collection of all its subsets, including the empty set and S itself.[1] The nth power set of S is obtained by iteratively applying the power set operation n times to S.[2]

Definition 1.1 (Base Set)**.** A *base set* S is the underlying set from which more elaborate structures, such as powersets and hyperstructures, are constructed. It is defined as

$$S = \{x \mid x \text{ belongs to a specified domain}\}.$$

All elements used in constructions like $\mathcal{P}(S)$ or $P_n(S)$ are drawn from the base set S.

Definition 1.2 (Powerset)**.** [1,3] The *powerset* of a set S, denoted by $\mathcal{P}(S)$, is the collection of all subsets of S, including both the empty set and S itself:

$$\mathcal{P}(S) = \{A \mid A \subseteq S\}.$$

DOI: https://doi.org/10.54216/PMTCS.060101

Received: Received: July 29, 2025 Revised: October 12, 2025 Accepted: December 27, 2025

Definition 1.3 (n-th Powerset)**.** (cf.[2])

The n-th powerset of a set H, denoted $\mathcal{P}_n(H)$, is defined iteratively, starting with the standard powerset. The recursive construction is given by:

$$\mathcal{P}_1(H) = \mathcal{P}(H), \quad \mathcal{P}_{n+1}(H) = \mathcal{P}(\mathcal{P}_n(H)), \quad \text{for } n \geq 1.$$

Similarly, the n-th non-empty powerset, denoted $\mathcal{P}_n^*(H)$, is defined recursively as:

$$\mathcal{P}_1^*(H) = \mathcal{P}^*(H), \quad \mathcal{P}_{n+1}^*(H) = \mathcal{P}^*(\mathcal{P}_n^*(H)).$$

Here, $\mathcal{P}^*(H)$ represents the powerset of H with the empty set removed.

Example 1.4 (Ensemble of Regression Models via the 2nd Powerset)**.** Let $H = \{X_1, X_2, X_3\}$ be a set of three predictor variables in a regression study. Then the first powerset

$$\mathcal{P}_1(H) = \mathcal{P}(H) = \big\{\emptyset,\ \{X_1\},\ \{X_2\},\ \{X_3\},\ \{X_1, X_2\},\ \{X_1, X_3\},\ \{X_2, X_3\},\ \{X_1, X_2, X_3\}\big\}$$

enumerates all possible subsets of predictors, each corresponding to a candidate regression model. Applying the powerset operation again yields the second powerset

$$\mathcal{P}_2(H) = \mathcal{P}\big(\mathcal{P}_1(H)\big),$$

whose elements are *sets of candidate models*. For instance,

$$U = \big\{\ \{X_1, X_2\},\ \{X_2, X_3\}\big\}\ \in\ \mathcal{P}_2(H)$$

represents an ensemble composed of the two regression models that use $\{X_1, X_2\}$ and $\{X_2, X_3\}$ respectively. This construction formalizes how one can organize and compare collections of models in statistical practice.

1.2 HyperProbability

Standard probability theory assigns each event a single real value in $[0, 1]$, satisfying non-negativity, normalization, and countable additivity (cf.[4,5]). HyperProbability assigns a set of probability values to each event, capturing uncertainty from multiple sources.[6] The definitions and examples of HyperProbability are presented below.

Definition 1.5 (Probability Space)**.** (cf.[7,8]) A *probability space* is a triple $(\Omega, \mathcal{F}, P)$ where

- Ω is a nonempty set called the *sample space*;
- $\mathcal{F}$ is a σ-algebra of subsets of Ω, called the *event space*;
- $P : \mathcal{F} \to [0, 1]$ is a function, called a *probability measure*, satisfying:
  1. $P(A) \geq 0$ for all $A \in \mathcal{F}$ (non-negativity),
  2. $P(\Omega) = 1$ (normalization),
  3. For any countable sequence of pairwise disjoint events $A_1, A_2, \dots \in \mathcal{F}$,

$$P\Big(\bigcup_{i=1}^{\infty} A_i\Big) = \sum_{i=1}^{\infty} P(A_i) \quad \text{(countable additivity).}$$

Definition 1.6 (Hyper-Probability)**.** [6] Let $(\Omega, \mathcal{F})$ be a measurable space, and let

$$\{p_k : \mathcal{F} \to [0, 1]\}_{k \in K}$$

be a nonempty family of probability measures on $(\Omega, \mathcal{F})$, where K is a nonempty index set.

DOI: https://doi.org/10.54216/PMTCS.060101
Received: Received: July 29, 2025 Revised: October 12, 2025 Accepted: December 27, 2025

The *Hyper-Probability* induced by $\{p_k\}_{k\in K}$ is the set-valued map

$$\mathrm{HP} : \mathcal{F} \longrightarrow \mathcal{P}([0,1])$$

defined by

$$\mathrm{HP}(A) := \{\, p_k(A) \mid k \in K \,\}, \qquad A \in \mathcal{F}.$$

Thus, for each event A, the value $\mathrm{HP}(A)$ collects all probability assessments assigned to A by the family of classical probability measures $\{p_k\}_{k\in K}$.

Example 1.7 (Hyper-Probability for a Tokyo rain forecast)**.** Let

$$\Omega = \{\text{Rain}, \text{No Rain}\}, \qquad \mathcal{F} = 2^{\Omega}.$$

Consider the event

$$A = \{\text{Rain}\}.$$

Suppose that three forecasting sources provide probability measures

$$p_{\text{JMA}}, \quad p_{\text{W}}, \quad p_{\text{App}} : \mathcal{F} \to [0,1],$$

such that

$$p_{\text{JMA}}(A) = 0.32, \qquad p_{\text{W}}(A) = 0.28, \qquad p_{\text{App}}(A) = 0.30.$$

Since each p_k is a probability measure, it is understood that

$$p_k(\varnothing) = 0, \qquad p_k(\Omega) = 1, \qquad p_k(\{\text{No Rain}\}) = 1 - p_k(A)$$

for $k \in \{\text{JMA}, \text{W}, \text{App}\}$.

Hence the induced Hyper-Probability satisfies

$$\mathrm{HP}(A) = \{\, p_{\text{JMA}}(A),\, p_{\text{W}}(A),\, p_{\text{App}}(A) \,\} = \{0.28,\ 0.30,\ 0.32\}.$$

Therefore, the event A is represented not by a single probability value, but by the set of multiple probability assessments $\{0.28, 0.30, 0.32\}$.

1.3 n-SuperHyperProbability

n-SuperHyperProbability extends HyperProbability by iterating the powerset construction, thereby representing hierarchical layers of uncertainty or belief aggregation.

To avoid ambiguity, we first introduce the iterated powerset notation.

Definition 1.8 (Iterated powerset)**.** Let X be a set. Define recursively

$$\mathcal{P}^{(0)}(X) := X, \qquad \mathcal{P}^{(n+1)}(X) := \mathcal{P}\big(\mathcal{P}^{(n)}(X)\big) \qquad (n \geq 0),$$

where $\mathcal{P}(X)$ denotes the powerset of X.

The definition of n-SuperHyperProbability is now given as follows.

DOI: https://doi.org/10.54216/PMTCS.060101
Received: Received: July 29, 2025 Revised: October 12, 2025 Accepted: December 27, 2025

Definition 1.9 (n-SuperHyperProbability)**.** [6] Let $(\Omega, \mathcal{F})$ be a measurable space, and let

$$\mathrm{HP} : \mathcal{F} \longrightarrow \mathcal{P}([0,1])$$

be a HyperProbability.

For each integer $n \geq 1$, the associated *n-SuperHyperProbability* is the map

$$\mathrm{SHP}^{(n)} : \mathcal{F} \longrightarrow \mathcal{P}^{(n)}([0,1])$$

defined recursively by

$$\mathrm{SHP}^{(1)}(A) := \mathrm{HP}(A),$$

and, for $n \geq 2$,

$$\mathrm{SHP}^{(n)}(A) := \mathcal{P}\big(\mathrm{SHP}^{(n-1)}(A)\big), \qquad A \in \mathcal{F}.$$

Equivalently,

$$\mathrm{SHP}^{(n)}(A) = \mathcal{P}^{(n-1)}\big(\mathrm{HP}(A)\big) \qquad (n \geq 1).$$

In particular, $\mathrm{SHP}^{(1)}$ coincides with HyperProbability.

Remark 1.10. If one wishes to include the classical case as level 0, one may adopt the convention

$$\mathcal{P}^{(0)}([0,1]) = [0,1]$$

and regard an ordinary probability measure

$$p : \mathcal{F} \to [0,1]$$

as a 0-level probability object. However, this level-0 object is not determined by HP alone in general, so it is better treated separately from Definition 1.9.

Example 1.11 (2-SuperHyperProbability in a Tokyo rain forecast)**.** Let

$$\Omega = \{\mathrm{Rain}, \mathrm{No\ Rain}\},$$

and let $\mathcal{F} = 2^{\Omega}$ be the corresponding σ-algebra. Consider the event

$$A = \{\mathrm{Rain}\}.$$

Assume that three forecasting sources provide the HyperProbability

$$\mathrm{HP}(A) = \{0.28,\ 0.30,\ 0.32\}.$$

Then, by Definition 1.9,

$$\mathrm{SHP}^{(2)}(A) = \mathcal{P}\big(\mathrm{HP}(A)\big) = \mathcal{P}\big(\{0.28, 0.30, 0.32\}\big).$$

Hence

$$\mathrm{SHP}^{(2)}(A) = \Big\{\varnothing,\ \{0.28\},\ \{0.30\},\ \{0.32\},\ \{0.28, 0.30\},\ \{0.28, 0.32\},\ \{0.30, 0.32\},\ \{0.28, 0.30, 0.32\}\Big\}.$$

Thus, a 2-SuperHyperProbability records not only individual probability assessments, but also all possible subcollections of those assessments, representing a second-order layer of uncertainty aggregation.

2 Main Results

This section presents the main results obtained in this paper.

Received: Received: July 29, 2025 Revised: October 12, 2025 Accepted: December 27, 2025

2.1 (m,n)-SuperHyperProbability

An (m,n)-SuperHyperProbability extends HyperProbability in two independent directions: the horizontal index m records iterated collections of events, while the vertical index n records iterated powerset levels of probability values.

For convenience, recall the iterated powerset notation:

$$\mathcal{P}^{(0)}(X) := X, \qquad \mathcal{P}^{(k+1)}(X) := \mathcal{P}\big(\mathcal{P}^{(k)}(X)\big) \qquad (k \geq 0).$$

Definition 2.1 ((m,n)-SuperHyperProbability)**.** Let $(\Omega, \mathcal{F})$ be a measurable space, and let

$$\mathrm{HP} : \mathcal{F} \longrightarrow \mathcal{P}([0,1])$$

be a Hyper-Probability.

For integers $m \geq 0$ and $n \geq 1$, define the map

$$\mathrm{SHP}^{(m,n)} : \mathcal{P}^{(m)}(\mathcal{F}) \longrightarrow \mathcal{P}^{(n)}([0,1])$$

recursively as follows.

(1) **Vertical base and vertical recursion ($m = 0$).** For every $A \in \mathcal{F}$,

$$\mathrm{SHP}^{(0,1)}(A) := \mathrm{HP}(A),$$

and for $n \geq 2$,

$$\mathrm{SHP}^{(0,n)}(A) := \mathcal{P}\big(\mathrm{SHP}^{(0,n-1)}(A)\big).$$

(2) **Horizontal recursion ($m \geq 1$).** For every $U \in \mathcal{P}^{(m)}(\mathcal{F})$, define

$$\mathrm{SHP}^{(m,n)}(U) := \bigcup_{A \in U} \mathrm{SHP}^{(m-1,n)}(A).$$

Remark 2.2. The classical single-valued probability level is best treated separately. If $P : \mathcal{F} \to [0,1]$ is an ordinary probability measure, one may regard P as the level $(0,0)$ object. However, for $m \geq 1$, the horizontal recursion is naturally formulated only for $n \geq 1$, because it involves unions of set-valued outputs.

Example 2.3 ($(1,2)$-SuperHyperProbability for Tokyo weather events)**.** Let

$$\Omega = \{\omega_{rs}, \omega_{r\bar{s}}, \omega_{\bar{r}s}, \omega_{\bar{r}\bar{s}}\},$$

where

ω_{rs} = rain and storm, $\omega_{r\bar{s}}$ = rain and no storm, $\omega_{\bar{r}s}$ = no rain and storm, $\omega_{\bar{r}\bar{s}}$ = no rain and no storm,

and let $\mathcal{F} = 2^{\Omega}$.

Define the events

$$A := \{\omega_{rs}, \omega_{r\bar{s}}\} \qquad \text{(rain)},$$
$$B := \{\omega_{rs}, \omega_{\bar{r}s}\} \qquad \text{(storm)}.$$

Suppose that

$$\mathrm{HP}(A) = \{0.28,\ 0.30,\ 0.32\}, \qquad \mathrm{HP}(B) = \{0.05,\ 0.08\}.$$

Let

$$U := \{A, B\} \in \mathcal{P}(\mathcal{F}).$$

Then

$$\mathrm{SHP}^{(1,2)}(U) = \mathrm{SHP}^{(0,2)}(A) \cup \mathrm{SHP}^{(0,2)}(B) = \mathcal{P}\big(\mathrm{HP}(A)\big) \cup \mathcal{P}\big(\mathrm{HP}(B)\big).$$

Hence

$$\mathrm{SHP}^{(1,2)}(U) = \mathcal{P}\big(\{0.28, 0.30, 0.32\}\big) \ \cup\ \mathcal{P}\big(\{0.05, 0.08\}\big).$$

Thus $\mathrm{SHP}^{(1,2)}(U)$ consists of all subsets of $\{0.28, 0.30, 0.32\}$ together with all subsets of $\{0.05, 0.08\}$, representing a second-order uncertainty structure over the two weather events.

DOI: https://doi.org/10.54216/PMTCS.060101

Received: Received: July 29, 2025 Revised: October 12, 2025 Accepted: December 27, 2025

Example 2.4 ($(2,1)$-SuperHyperProbability for aggregated weather portfolios)**.** Using the same measurable space and the same events A and B, define

$$U_1 := \{A, B\}, \qquad U_2 := \{A\}, \qquad V := \{U_1, U_2\} \in \mathcal{P}^{(2)}(\mathcal{F}).$$

Then

$$\mathrm{SHP}^{(2,1)}(V) = \mathrm{SHP}^{(1,1)}(U_1) \cup \mathrm{SHP}^{(1,1)}(U_2).$$

Now

$$\mathrm{SHP}^{(1,1)}(U_1) = \mathrm{HP}(A) \cup \mathrm{HP}(B), \qquad \mathrm{SHP}^{(1,1)}(U_2) = \mathrm{HP}(A),$$

so

$$\mathrm{SHP}^{(2,1)}(V) = \big(\mathrm{HP}(A) \cup \mathrm{HP}(B)\big) \cup \mathrm{HP}(A) = \{0.05,\ 0.08,\ 0.28,\ 0.30,\ 0.32\}.$$

This value aggregates all probability assessments arising from the two first-level weather portfolios.

Example 2.5 (Supply-chain failure risk as a $(2,2)$-SuperHyperProbability)**.** Let

$$\Omega = \{\omega_{00}, \omega_{10}, \omega_{01}, \omega_{11}\},$$

where ω_{ij} indicates whether supplier X fails ($i = 1$) or not ($i = 0$), and whether supplier Y fails ($j = 1$) or not ($j = 0$). Let $\mathcal{F} = 2^{\Omega}$, and define

$$A := \{\omega_{10}, \omega_{11}\} \qquad \text{(supplier } X \text{ fails)},$$

$$B := \{\omega_{01}, \omega_{11}\} \qquad \text{(supplier } Y \text{ fails)}.$$

Assume that

$$\mathrm{HP}(A) = \{0.20,\ 0.30\}, \qquad \mathrm{HP}(B) = \{0.50,\ 0.70\}.$$

Set

$$U_A := \{A\}, \qquad U_B := \{B\}, \qquad V := \{U_A, U_B\} \in \mathcal{P}^{(2)}(\mathcal{F}).$$

Then

$$\mathrm{SHP}^{(2,2)}(V) = \mathrm{SHP}^{(1,2)}(U_A) \cup \mathrm{SHP}^{(1,2)}(U_B) = \mathcal{P}\big(\mathrm{HP}(A)\big) \cup \mathcal{P}\big(\mathrm{HP}(B)\big).$$

Therefore,

$$\mathrm{SHP}^{(2,2)}(V) = \mathcal{P}\big(\{0.20, 0.30\}\big) \cup \mathcal{P}\big(\{0.50, 0.70\}\big),$$

that is,

$$\mathrm{SHP}^{(2,2)}(V) = \Big\{\varnothing,\ \{0.20\},\ \{0.30\},\ \{0.20, 0.30\},\ \{0.50\},\ \{0.70\},\ \{0.50, 0.70\}\Big\}.$$

This gives a second-order, two-portfolio uncertainty representation for supplier-failure risk.

Example 2.6 (Election uncertainty as a $(1,3)$-SuperHyperProbability)**.** Let

$$\Omega = \{\mathrm{Win}, \mathrm{Lose}\}, \qquad \mathcal{F} = 2^{\Omega}, \qquad E := \{\mathrm{Win}\}.$$

Suppose that polling sources provide

$$\mathrm{HP}(E) = \{0.42,\ 0.45,\ 0.47\}.$$

Let

$$U := \{E\} \in \mathcal{P}(\mathcal{F}).$$

Then

$$\mathrm{SHP}^{(1,3)}(U) = \mathrm{SHP}^{(0,3)}(E) = \mathcal{P}\Big(\mathcal{P}\big(\mathrm{HP}(E)\big)\Big).$$

Hence

$$\mathrm{SHP}^{(1,3)}(U) = \mathcal{P}\Big(\mathcal{P}\big(\{0.42, 0.45, 0.47\}\big)\Big).$$

Since $\mathcal{P}(\{0.42, 0.45, 0.47\})$ has 8 elements, the set $\mathrm{SHP}^{(1,3)}(U)$ has $2^8 = 256$ elements. For example,

$$\big\{\{0.42\}, \{0.45, 0.47\}\big\} \in \mathrm{SHP}^{(1,3)}(U).$$

This is a third-order uncertainty object built from the underlying election-win assessments.

Theorem 2.7 (Well-definedness of (m, n)-SuperHyperProbability)**.** *For every $m \geq 0$ and $n \geq 1$, the map*

$$\mathrm{SHP}^{(m,n)} : \mathcal{P}^{(m)}(\mathcal{F}) \to \mathcal{P}^{(n)}([0,1])$$

in Definition 2.1 is well-defined.

Moreover, for every $n \geq 1$,

$$\mathrm{SHP}^{(0,n)} = \mathrm{SHP}^{(n)},$$

where $\mathrm{SHP}^{(n)}$ denotes the usual n-SuperHyperProbability induced by HP.

Proof. Fix $n \geq 1$. We prove by induction on m that

$$\mathrm{SHP}^{(m,n)}(U) \in \mathcal{P}^{(n)}([0,1]) \qquad \text{for all } U \in \mathcal{P}^{(m)}(\mathcal{F}).$$

Base case: $m = 0$. If $m = 0$, then $U = A \in \mathcal{F}$. For $n = 1$,

$$\mathrm{SHP}^{(0,1)}(A) = \mathrm{HP}(A) \in \mathcal{P}([0,1]) = \mathcal{P}^{(1)}([0,1]).$$

For $n \geq 2$, assume inductively that

$$\mathrm{SHP}^{(0,n-1)}(A) \in \mathcal{P}^{(n-1)}([0,1]).$$

Then

$$\mathrm{SHP}^{(0,n)}(A) = \mathcal{P}\big(\mathrm{SHP}^{(0,n-1)}(A)\big) \in \mathcal{P}\big(\mathcal{P}^{(n-1)}([0,1])\big) = \mathcal{P}^{(n)}([0,1]).$$

Hence $\mathrm{SHP}^{(0,n)}$ is well-defined for every $n \geq 1$.

Inductive step on m. Assume that for some $m \geq 0$, the map

$$\mathrm{SHP}^{(m,n)} : \mathcal{P}^{(m)}(\mathcal{F}) \to \mathcal{P}^{(n)}([0,1])$$

is well-defined. Let $U \in \mathcal{P}^{(m+1)}(\mathcal{F})$. Then U is a subset of $\mathcal{P}^{(m)}(\mathcal{F})$, and for each $A \in U$,

$$\mathrm{SHP}^{(m,n)}(A) \in \mathcal{P}^{(n)}([0,1]).$$

Since $n \geq 1$, every element of $\mathcal{P}^{(n)}([0,1])$ is a subset of $\mathcal{P}^{(n-1)}([0,1])$. Therefore,

$$\bigcup_{A \in U} \mathrm{SHP}^{(m,n)}(A) \subseteq \mathcal{P}^{(n-1)}([0,1]),$$

and hence

$$\bigcup_{A \in U} \mathrm{SHP}^{(m,n)}(A) \in \mathcal{P}\big(\mathcal{P}^{(n-1)}([0,1])\big) = \mathcal{P}^{(n)}([0,1]).$$

Thus

$$\mathrm{SHP}^{(m+1,n)}(U) = \bigcup_{A \in U} \mathrm{SHP}^{(m,n)}(A)$$

is well-defined.

By induction on m, $\mathrm{SHP}^{(m,n)}$ is well-defined for all $m \geq 0$ and $n \geq 1$.

Finally, when $m = 0$, the above definition is exactly the recursive definition of the usual n-SuperHyperProbability induced by HP. Hence

$$\mathrm{SHP}^{(0,n)} = \mathrm{SHP}^{(n)} \qquad (n \geq 1).$$

This completes the proof. ☐

DOI: https://doi.org/10.54216/PMTCS.060101
Received: Received: July 29, 2025 Revised: October 12, 2025 Accepted: December 27, 2025

2.2 Hyper-Probability Measure

A probability measure is a map from a σ-algebra to the unit interval $[0, 1]$ satisfying the usual axioms of non-negativity, normalization, and countable additivity. In this subsection, we first recall the classical notion of probability measure, then introduce Hyper-Probability measures and their higher-order extension, namely (m, n)-SuperHyperProbability measures.

Definition 2.8 (Probability Measure)**.** (cf.[9]) Let $(\Omega, \mathcal{F})$ be a measurable space, where Ω is a nonempty set and $\mathcal{F}$ is a σ-algebra of subsets of Ω. A *probability measure* on $(\Omega, \mathcal{F})$ is a function

$$\mu : \mathcal{F} \longrightarrow [0, 1]$$

such that:

(1) **Non-negativity:**

$$\mu(E) \geq 0 \qquad \text{for all } E \in \mathcal{F}.$$

(2) **Null empty set:**

$$\mu(\varnothing) = 0.$$

(3) **Normalization:**

$$\mu(\Omega) = 1.$$

(4) **Countable additivity:** whenever $\{E_i\}_{i=1}^{\infty} \subseteq \mathcal{F}$ is a sequence of pairwise disjoint sets,

$$\mu\Big(\bigcup_{i=1}^{\infty} E_i\Big) = \sum_{i=1}^{\infty} \mu(E_i).$$

Example 2.9 (Fair six-sided die)**.** Let

$$\Omega = \{1, 2, 3, 4, 5, 6\}, \qquad \mathcal{F} = 2^{\Omega}.$$

Define

$$\mu(A) := \frac{|A|}{6} \qquad (A \subseteq \Omega).$$

Then μ is a probability measure on $(\Omega, \mathcal{F})$. For the event

$$E = \{2, 4, 6\},$$

one has

$$\mu(E) = \frac{3}{6} = \frac{1}{2}.$$

Example 2.10 (Binary diagnostic outcome)**.** Let

$$\Omega = \{\text{Positive}, \text{Negative}\}, \qquad \mathcal{F} = 2^{\Omega}.$$

Suppose

$$\mu(\{\text{Positive}\}) = 0.02, \qquad \mu(\{\text{Negative}\}) = 0.98.$$

Then μ extends uniquely to a probability measure on $(\Omega, \mathcal{F})$, and

$$\mu(\Omega) = 1.$$

This models a binary outcome with positive probability 0.02.

Example 2.11 (Poisson arrival counts)**.** Let

$$\Omega = \mathbb{N}_0 = \{0, 1, 2, \dots\}, \qquad \mathcal{F} = 2^{\mathbb{N}_0}.$$

Fix $\lambda = 4$, and define

$$\mu(\{k\}) = e^{-4}\frac{4^k}{k!}, \qquad k \in \mathbb{N}_0.$$

DOI: https://doi.org/10.54216/PMTCS.060101
Received: Received: July 29, 2025 Revised: October 12, 2025 Accepted: December 27, 2025

For any $A \subseteq \mathbb{N}_0$, set

$$\mu(A) := \sum_{k \in A} e^{-4} \frac{4^k}{k!}.$$

Then μ is a probability measure on $(\mathbb{N}_0, 2^{\mathbb{N}_0})$. For example,

$$\mu(\{0,1,2\}) = e^{-4}\Big(1 + 4 + \frac{4^2}{2!}\Big) \approx 0.2381.$$

Definition 2.12 (Hyper-Probability Measure)**.** Let $(\Omega, \mathcal{F})$ be a measurable space. A *Hyper-Probability measure* on $(\Omega, \mathcal{F})$ is a set-valued map

$$\mathrm{HP} : \mathcal{F} \longrightarrow \mathcal{P}([0,1])$$

for which there exist a nonempty index set K and a family of probability measures

$$\{\mu_k : \mathcal{F} \to [0,1]\}_{k \in K}$$

such that, for every $A \in \mathcal{F}$,

$$\mathrm{HP}(A) = \{\mu_k(A) \mid k \in K\}.$$

Remark 2.13. A Hyper-Probability measure is generally not a single probability measure, but rather a set-valued aggregation of probability values arising from a family of ordinary probability measures.

Example 2.14 (Weather forecast aggregation)**.** Let

$$\Omega = \{\text{Rain}, \text{No Rain}\}, \qquad \mathcal{F} = 2^{\Omega}, \qquad A = \{\text{Rain}\}.$$

Suppose three probability measures satisfy

$$\mu_{\text{JMA}}(A) = 0.32, \qquad \mu_{\text{Wthr}}(A) = 0.28, \qquad \mu_{\text{App}}(A) = 0.30.$$

Then

$$\mathrm{HP}(A) = \{0.28, 0.30, 0.32\}.$$

Example 2.15 (Medical diagnosis from multiple sources)**.** Let

$$\Omega = \{\text{Positive}, \text{Negative}\}, \qquad \mathcal{F} = 2^{\Omega}, \qquad B = \{\text{Positive}\}.$$

Suppose

$$\mu_{\text{Clin}}(B) = 0.07, \qquad \mu_{\text{ML}}(B) = 0.05, \qquad \mu_{\text{Lab}}(B) = 0.09.$$

Then

$$\mathrm{HP}(B) = \{0.05, 0.07, 0.09\}.$$

Example 2.16 (Credit default risk assessment)**.** Let

$$\Omega = \{\text{Default}, \text{No Default}\}, \qquad \mathcal{F} = 2^{\Omega}, \qquad C = \{\text{Default}\}.$$

Suppose

$$\mu_{\text{S\&P}}(C) = 0.02, \qquad \mu_{\text{Moody}}(C) = 0.03, \qquad \mu_{\text{Fitch}}(C) = 0.025.$$

Then

$$\mathrm{HP}(C) = \{0.02, 0.025, 0.03\}.$$

Theorem 2.17 (Singleton-valued recovery of classical probability)**.** *Let* HP *be a Hyper-Probability measure induced by a family* $\{\mu_k\}_{k \in K}$*. If* $K = \{k_0\}$ *is a singleton, then*

$$\mathrm{HP}(A) = \{\mu_{k_0}(A)\} \qquad \text{for all } A \in \mathcal{F}.$$

Hence HP *is the singleton-valued embedding of the ordinary probability measure* μ_{k_0}*.*

Proof. If $K = \{k_0\}$, then by Definition 2.12,

$$\mathrm{HP}(A) = \{\mu_{k_0}(A)\} \qquad (A \in \mathcal{F}).$$

Since μ_{k_0} is a probability measure, the claim follows immediately. □

Theorem 2.18 (Hyper-additivity and boundary values)**.** *Let* HP *be a Hyper-Probability measure induced by a family* $\{\mu_k\}_{k\in K}$*. Then:*

(a)
$$\mathrm{HP}(\varnothing) = \{0\}.$$

(b)
$$\mathrm{HP}(\Omega) = \{1\}.$$

(c) *If* $\{A_i\}_{i=1}^{\infty} \subseteq \mathcal{F}$ *is a sequence of pairwise disjoint sets, then*
$$\mathrm{HP}\Big(\bigcup_{i=1}^{\infty} A_i\Big) = \left\{ \sum_{i=1}^{\infty} \mu_k(A_i) \;\middle|\; k \in K \right\}.$$

Proof. For each $k \in K$, since μ_k is a probability measure,
$$\mu_k(\varnothing) = 0, \qquad \mu_k(\Omega) = 1,$$
and for pairwise disjoint $\{A_i\}_{i=1}^{\infty}$,
$$\mu_k\Big(\bigcup_{i=1}^{\infty} A_i\Big) = \sum_{i=1}^{\infty} \mu_k(A_i).$$
Taking the set of these values over all $k \in K$ yields
$$\mathrm{HP}(\varnothing) = \{0\}, \qquad \mathrm{HP}(\Omega) = \{1\},$$
and
$$\mathrm{HP}\Big(\bigcup_{i=1}^{\infty} A_i\Big) = \left\{ \sum_{i=1}^{\infty} \mu_k(A_i) \mid k \in K \right\}.$$
□

We now introduce the higher-order extension of Hyper-Probability measures.

Definition 2.19 (Iterated powerset)**.** Let X be a set. Define recursively
$$\mathcal{P}^{(0)}(X) := X, \qquad \mathcal{P}^{(n+1)}(X) := \mathcal{P}\big(\mathcal{P}^{(n)}(X)\big) \qquad (n \geq 0).$$

Definition 2.20 ((m,n)-SuperHyperProbability Measure)**.** Let $(\Omega, \mathcal{F})$ be a measurable space, and let
$$\{\mu_k : \mathcal{F} \to [0,1]\}_{k\in K}$$
be a nonempty family of probability measures on $(\Omega, \mathcal{F})$. Let HP be the induced Hyper-Probability measure,
$$\mathrm{HP}(A) = \{\mu_k(A) \mid k \in K\}, \qquad A \in \mathcal{F}.$$

For integers $m \geq 0$ and $n \geq 1$, the associated (m,n)-*SuperHyperProbability measure* is the map
$$\mathrm{SHP}^{(m,n)} : \mathcal{P}^{(m)}(\mathcal{F}) \longrightarrow \mathcal{P}^{(n)}([0,1])$$
defined recursively by:

(1) **Vertical base and vertical recursion (**$m = 0$**).** For every $A \in \mathcal{F}$,
$$\mathrm{SHP}^{(0,1)}(A) := \mathrm{HP}(A),$$
and for $n \geq 2$,
$$\mathrm{SHP}^{(0,n)}(A) := \mathcal{P}\big(\mathrm{SHP}^{(0,n-1)}(A)\big).$$

DOI: https://doi.org/10.54216/PMTCS.060101
Received: Received: July 29, 2025 Revised: October 12, 2025 Accepted: December 27, 2025

(2) **Horizontal recursion ($m \geq 1$).** For every $U \in \mathcal{P}^{(m)}(\mathcal{F})$, define

$$\mathrm{SHP}^{(m,n)}(U) := \bigcup_{A \in U} \mathrm{SHP}^{(m-1,n)}(A).$$

Remark 2.21. The ordinary scalar-valued probability measure belongs to level 0 and is best treated separately. The construction in Definition 2.20 starts at the first set-valued level $n = 1$, namely the Hyper-Probability measure.

Example 2.22 (($2, 2$)-SuperHyperProbability in intrusion-alert assessment)**.** Let

$$\Omega = \{\omega_{sa}, \omega_{s\bar{a}}, \omega_{\bar{s}a}, \omega_{\bar{s}\bar{a}}\},$$

where

$$\omega_{sa} = \text{signature alert and anomaly alert}, \quad \omega_{s\bar{a}} = \text{signature alert and no anomaly alert},$$
$$\omega_{\bar{s}a} = \text{no signature alert and anomaly alert}, \quad \omega_{\bar{s}\bar{a}} = \text{no signature alert and no anomaly alert}.$$

Let $\mathcal{F} = 2^{\Omega}$, and define the events

$$A = \{\omega_{sa}, \omega_{s\bar{a}}\}, \qquad B = \{\omega_{sa}, \omega_{\bar{s}a}\}.$$

Assume that

$$\mu_1(A) = 0.70, \quad \mu_2(A) = 0.65, \quad \mu_3(A) = 0.75,$$
$$\mu_1(B) = 0.60, \quad \mu_2(B) = 0.55, \quad \mu_3(B) = 0.65.$$

Then

$$\mathrm{HP}(A) = \{0.65, 0.70, 0.75\}, \qquad \mathrm{HP}(B) = \{0.55, 0.60, 0.65\}.$$

Define

$$U_A = \{A\}, \qquad U_B = \{B\}, \qquad V = \{U_A, U_B\} \in \mathcal{P}^{(2)}(\mathcal{F}).$$

Then

$$\mathrm{SHP}^{(2,2)}(V) = \mathrm{SHP}^{(1,2)}(U_A) \cup \mathrm{SHP}^{(1,2)}(U_B) = \mathcal{P}(\mathrm{HP}(A)) \cup \mathcal{P}(\mathrm{HP}(B)).$$

That is,

$$\mathrm{SHP}^{(2,2)}(V) = \mathcal{P}\big(\{0.65, 0.70, 0.75\}\big) \cup \mathcal{P}\big(\{0.55, 0.60, 0.65\}\big).$$

Example 2.23 (($1, 3$)-SuperHyperProbability in credit default forecasting)**.** Let

$$\Omega = \{\mathrm{Default}, \mathrm{No\ Default}\}, \qquad \mathcal{F} = 2^{\Omega}, \qquad C = \{\mathrm{Default}\}.$$

Suppose

$$\mu_A(C) = 0.02, \qquad \mu_B(C) = 0.03, \qquad \mu_C(C) = 0.025.$$

Then

$$\mathrm{HP}(C) = \{0.02, 0.025, 0.03\}.$$

Let

$$U = \{C\} \in \mathcal{P}(\mathcal{F}).$$

Then

$$\mathrm{SHP}^{(1,3)}(U) = \mathrm{SHP}^{(0,3)}(C) = \mathcal{P}\Big(\mathcal{P}\big(\{0.02, 0.025, 0.03\}\big)\Big).$$

Since $\mathcal{P}(\{0.02, 0.025, 0.03\})$ has 8 elements, the set

$$\mathrm{SHP}^{(1,3)}(U)$$

has $2^8 = 256$ elements. For example,

$$\big\{\{0.02, 0.025\}, \{0.025, 0.03\}\big\} \in \mathrm{SHP}^{(1,3)}(U).$$

Theorem 2.24 (Specialization to Hyper-Probability and singleton-valued classical probability)**.** *Under Definition 2.20, the following hold:*

(a) *For every $A \in \mathcal{F}$,*

$$\mathrm{SHP}^{(0,1)}(A) = \mathrm{HP}(A).$$

DOI: https://doi.org/10.54216/PMTCS.060101
Received: Received: July 29, 2025 Revised: October 12, 2025 Accepted: December 27, 2025

(b) *If* $K = \{k_0\}$ *is a singleton, then*

$$\mathrm{SHP}^{(0,1)}(A) = \{\mu_{k_0}(A)\} \qquad \textit{for all } A \in \mathcal{F}.$$

Hence $\mathrm{SHP}^{(0,1)}$ *is the singleton-valued embedding of the ordinary probability measure* μ_{k_0}*.*

Proof. Part (a) is immediate from the definition. For part (b), if $K = \{k_0\}$, then

$$\mathrm{SHP}^{(0,1)}(A) = \mathrm{HP}(A) = \{\mu_{k_0}(A)\} \qquad (A \in \mathcal{F}).$$

This is exactly the singleton-valued representation of the classical probability measure μ_{k_0}. □

Theorem 2.25 (Well-definedness of (m, n)-SuperHyperProbability measures)**.** *For every* $m \geq 0$ *and* $n \geq 1$*, the map*

$$\mathrm{SHP}^{(m,n)} : \mathcal{P}^{(m)}(\mathcal{F}) \longrightarrow \mathcal{P}^{(n)}([0,1])$$

given in Definition 2.20 is well-defined.

Proof. Fix $n \geq 1$. We prove by induction on m that

$$\mathrm{SHP}^{(m,n)}(U) \in \mathcal{P}^{(n)}([0,1]) \qquad \text{for every } U \in \mathcal{P}^{(m)}(\mathcal{F}).$$

Base case: $m = 0$**.** Let $A \in \mathcal{F}$. For $n = 1$,

$$\mathrm{SHP}^{(0,1)}(A) = \mathrm{HP}(A) \in \mathcal{P}([0,1]) = \mathcal{P}^{(1)}([0,1]).$$

For $n \geq 2$, assuming inductively that

$$\mathrm{SHP}^{(0,n-1)}(A) \in \mathcal{P}^{(n-1)}([0,1]),$$

we obtain

$$\mathrm{SHP}^{(0,n)}(A) = \mathcal{P}\big(\mathrm{SHP}^{(0,n-1)}(A)\big) \in \mathcal{P}\big(\mathcal{P}^{(n-1)}([0,1])\big) = \mathcal{P}^{(n)}([0,1]).$$

Inductive step on m**.** Assume that

$$\mathrm{SHP}^{(m,n)} : \mathcal{P}^{(m)}(\mathcal{F}) \to \mathcal{P}^{(n)}([0,1])$$

is well-defined. Let $U \in \mathcal{P}^{(m+1)}(\mathcal{F})$. Then $U \subseteq \mathcal{P}^{(m)}(\mathcal{F})$, so for every $A \in U$,

$$\mathrm{SHP}^{(m,n)}(A) \in \mathcal{P}^{(n)}([0,1]).$$

Since $n \geq 1$, every element of $\mathcal{P}^{(n)}([0,1])$ is a subset of $\mathcal{P}^{(n-1)}([0,1])$. Hence

$$\bigcup_{A \in U} \mathrm{SHP}^{(m,n)}(A) \subseteq \mathcal{P}^{(n-1)}([0,1]),$$

and therefore

$$\bigcup_{A \in U} \mathrm{SHP}^{(m,n)}(A) \in \mathcal{P}\big(\mathcal{P}^{(n-1)}([0,1])\big) = \mathcal{P}^{(n)}([0,1]).$$

Thus

$$\mathrm{SHP}^{(m+1,n)}(U) = \bigcup_{A \in U} \mathrm{SHP}^{(m,n)}(A)$$

is well-defined. □

Theorem 2.26 (Structural Properties of (m,n)-SHP)**.** *For $n \geq 1$, define recursively*

$$Z_1 := \{0\}, \qquad Z_n := \mathcal{P}(Z_{n-1}) \quad (n \geq 2),$$

and

$$O_1 := \{1\}, \qquad O_n := \mathcal{P}(O_{n-1}) \quad (n \geq 2).$$

Also define the m-fold singleton chain of Ω by

$$\Omega^{(0)} := \Omega, \qquad \Omega^{(m+1)} := \{\Omega^{(m)}\} \quad (m \geq 0).$$

Then the (m,n)-SuperHyperProbability measure $\mathrm{SHP}^{(m,n)}$ satisfies:

(a) *For $m = 0$,*

$$\mathrm{SHP}^{(0,n)}(\varnothing) = Z_n.$$

For every $m \geq 1$,

$$\mathrm{SHP}^{(m,n)}(\varnothing) = \varnothing.$$

(b) *For every $m \geq 0$ and $n \geq 1$,*

$$\mathrm{SHP}^{(m,n)}\big(\Omega^{(m)}\big) = O_n.$$

(c) *Let $m \geq 1$. For any family $\{U_j\}_{j\in J}$ with*

$$U_j \subseteq \mathcal{P}^{(m-1)}(\mathcal{F}) \qquad (j \in J),$$

equivalently $U_j \in \mathcal{P}^{(m)}(\mathcal{F})$, one has

$$\mathrm{SHP}^{(m,n)}\Big(\bigcup_{j\in J} U_j\Big) = \bigcup_{j\in J} \mathrm{SHP}^{(m,n)}(U_j).$$

Proof. (a) First let $m = 0$. Since every μ_k is a probability measure,

$$\mathrm{SHP}^{(0,1)}(\varnothing) = \mathrm{HP}(\varnothing) = \{0\} = Z_1.$$

For $n \geq 2$, by vertical recursion,

$$\mathrm{SHP}^{(0,n)}(\varnothing) = \mathcal{P}\big(\mathrm{SHP}^{(0,n-1)}(\varnothing)\big) = \mathcal{P}(Z_{n-1}) = Z_n.$$

Hence $\mathrm{SHP}^{(0,n)}(\varnothing) = Z_n$ for all $n \geq 1$.

Now let $m \geq 1$. Then

$$\mathrm{SHP}^{(m,n)}(\varnothing) = \bigcup_{A\in\varnothing} \mathrm{SHP}^{(m-1,n)}(A) = \varnothing.$$

(b) For $m \geq 1$, since $\Omega^{(m)} = \{\Omega^{(m-1)}\}$, we have

$$\mathrm{SHP}^{(m,n)}\big(\Omega^{(m)}\big) = \bigcup_{A\in\Omega^{(m)}} \mathrm{SHP}^{(m-1,n)}(A) = \mathrm{SHP}^{(m-1,n)}\big(\Omega^{(m-1)}\big).$$

Iterating yields

$$\mathrm{SHP}^{(m,n)}\big(\Omega^{(m)}\big) = \mathrm{SHP}^{(0,n)}(\Omega).$$

Since each $\mu_k(\Omega) = 1$,

$$\mathrm{SHP}^{(0,1)}(\Omega) = \mathrm{HP}(\Omega) = \{1\} = O_1.$$

For $n \geq 2$,

$$\mathrm{SHP}^{(0,n)}(\Omega) = \mathcal{P}\big(\mathrm{SHP}^{(0,n-1)}(\Omega)\big) = \mathcal{P}(O_{n-1}) = O_n.$$

Therefore,

$$\mathrm{SHP}^{(m,n)}\big(\Omega^{(m)}\big) = O_n$$

Received: Received: July 29, 2025 Revised: October 12, 2025 Accepted: December 27, 2025

for all $m \geq 0$ and $n \geq 1$.

(c) Let $m \geq 1$. Then

$$\mathrm{SHP}^{(m,n)}\Big(\bigcup_{j\in J} U_j\Big) = \bigcup_{A\in\bigcup_{j\in J} U_j} \mathrm{SHP}^{(m-1,n)}(A).$$

Using

$$\bigcup_{A\in\bigcup_{j\in J} U_j} X_A = \bigcup_{j\in J}\bigcup_{A\in U_j} X_A,$$

with $X_A := \mathrm{SHP}^{(m-1,n)}(A)$, we obtain

$$\mathrm{SHP}^{(m,n)}\Big(\bigcup_{j\in J} U_j\Big) = \bigcup_{j\in J}\bigcup_{A\in U_j} \mathrm{SHP}^{(m-1,n)}(A) = \bigcup_{j\in J} \mathrm{SHP}^{(m,n)}(U_j).$$

This proves the claim. □

2.3 Hyper-Probability Distribution

A probability distribution describes how probabilities are assigned to values or events of a random variable across its sample space(cf.[10–12]). In this subsection, we extend the classical probability distribution concept by introducing the frameworks of Hyper-Probability and Super-Hyper-Probability.

Definition 2.27 (Probability Distribution)**.** (cf.[13–15]) Let $(\Omega, \mathcal{F}, P)$ be a probability space and let $(E, \mathcal{E})$ be a measurable space. A *random variable* $X\colon \Omega \to E$ is a measurable function, i.e. $X^{-1}(B) \in \mathcal{F}$ for every $B \in \mathcal{E}$. The *probability distribution* (or *law*) of X is the pushforward measure

$$\mathcal{L}_X = X_*P\colon \mathcal{E} \longrightarrow [0,1], \quad X_*P(B) = P\big(X^{-1}(B)\big), \quad B \in \mathcal{E}.$$

This single definition encompasses both discrete and continuous cases:

(1) **Discrete distributions.** If E is countable (or finite), then X has a *probability mass function* (pmf)

$$p_X(x) = P\big(X = x\big), \quad x \in E,$$

satisfying $\sum_{x\in E} p_X(x) = 1$ and $\mathcal{L}_X(\{x\}) = p_X(x)$.

(2) **Absolutely continuous distributions.** If $E = \mathbb{R}$ (or $\mathbb{R}^d$) and $\mathcal{L}_X \ll \lambda$ (Lebesgue measure), then there exists a *probability density function* (pdf)

$$f_X(x) \geq 0, \quad \int_E f_X(x)\,dx = 1, \quad \mathcal{L}_X(B) = \int_B f_X(x)\,dx \quad (\forall\, B \in \mathcal{B}(E)).$$

(3) **Cumulative distribution function (CDF).** When $E = \mathbb{R}$, the *cdf* of X is

$$F_X(x) = \mathcal{L}_X\big((-\infty, x]\big) = P\big(X \leq x\big), \quad x \in \mathbb{R},$$

which satisfies:

- non-decreasing: $x_1 < x_2 \implies F_X(x_1) \leq F_X(x_2)$,
- right-continuous: $\lim_{h\downarrow 0} F_X(x+h) = F_X(x)$,
- limits: $\lim_{x\to-\infty} F_X(x) = 0,\ \lim_{x\to\infty} F_X(x) = 1$.

Example 2.28 (Discrete Uniform Distribution: Fair Six-Sided Die)**.** Let $\Omega = \{1,2,3,4,5,6\}$, $\mathcal{F} = 2^\Omega$, and let $X\colon \Omega \to \Omega$ be the identity map representing the outcome of a die roll. Then X has the discrete uniform distribution with probability mass function

$$p_X(k) = P\big(X = k\big) = \frac{1}{6}, \quad k = 1, 2, \ldots, 6,$$

so that for any event $A \subseteq \Omega$,

$$\mathcal{L}_X(A) = \sum_{k\in A} p_X(k) = \frac{|A|}{6}.$$

For example, the probability of rolling an even number is $\mathcal{L}_X(\{2,4,6\}) = 3/6 = 0.5$.

DOI: https://doi.org/10.54216/PMTCS.060101

Received: Received: July 29, 2025 Revised: October 12, 2025 Accepted: December 27, 2025

Example 2.29 (Poisson Distribution: Customer Arrivals)**.** Let $\Omega = \{0, 1, 2, \dots\}$, $\mathcal{F} = 2^{\Omega}$, and let X be the number of customers entering a café in one hour. Historical data suggest a Poisson model with mean $\lambda = 5$. Then

$$p_X(k) = P\big(X = k\big) = e^{-5}\,\frac{5^k}{k!}, \quad k = 0, 1, 2, \dots,$$

and for any $A \subseteq \Omega$,

$$\mathcal{L}_X(A) = \sum_{k \in A} e^{-5}\,\frac{5^k}{k!}.$$

For instance, the probability of at most 3 customers is $\sum_{k=0}^{3} e^{-5} 5^k / k! \approx 0.265$.

Example 2.30 (Normal Distribution: Daily High Temperature)**.** Let $\Omega = \mathbb{R}$, $\mathcal{F} = \mathcal{B}(\mathbb{R})$, and let X be the high temperature (°C) in a city on a given day. Meteorological records show $X \sim N(\mu = 20, \sigma^2 = 4)$. Its probability density function is

$$f_X(x) = \frac{1}{\sqrt{2\pi}\,2} \exp\Big(-\frac{(x-20)^2}{2 \cdot 4}\Big), \quad x \in \mathbb{R},$$

and for any Borel set $B \subseteq \mathbb{R}$,

$$\mathcal{L}_X(B) = \int_B \frac{1}{2\sqrt{2\pi}} e^{-\frac{(x-20)^2}{8}}\,dx.$$

For example, the probability of a "warm" day between 22 °C and 25 °C is $\int_{22}^{25} f_X(x)\,dx \approx 0.261$.

Definition 2.31 (Hyper-Probability Distribution)**.** Let $(\Omega, \mathcal{F}, P)$ be a probability space and $(E, \mathcal{E})$ a measurable space. Fix a nonempty index set K and a family of probability measures $\{P_k \colon \mathcal{F} \to [0,1]\}_{k \in K}$. A *hyper-random variable* is a measurable map

$$X \colon \Omega \longrightarrow E, \quad X^{-1}(B) \in \mathcal{F}\, \forall\, B \in \mathcal{E}.$$

Its *hyper-probability distribution* is the set-valued pushforward

$$\mathcal{HP}_X \;=\; X_{*}\{P_k\}_{k \in K} \;:\; \mathcal{E} \longrightarrow \mathcal{P}([0,1]),$$

defined by

$$\mathcal{HP}_X(B) \;=\; \big\{\, P_k\big(X^{-1}(B)\big) \;\big|\; k \in K \big\}, \quad B \in \mathcal{E}.$$

Thus $\mathcal{HP}_X(B)$ collects all classical probabilities that $X \in B$ under the different measures P_k.

Example 2.32 (Medical Diagnosis via Multiple Classifiers)**.** Let Ω be the population of patients presenting a particular symptom, and let $(E, \mathcal{E}) = \big\{\text{Positive}, \text{Negative}\big\}$ with the discrete σ-algebra. Define

$$X(\omega) = \begin{cases} \text{Positive}, & \text{if patient } \omega \text{ has the disease,} \\ \text{Negative}, & \text{otherwise.} \end{cases}$$

Three diagnostic procedures give:

$$P_{\text{Clin}}(X = \text{Positive}) = 0.07, \quad P_{\text{ML}}(X = \text{Positive}) = 0.05, \quad P_{\text{Lab}}(X = \text{Positive}) = 0.09.$$

Thus for the event $C = \{X = \text{Positive}\}$,

$$\mathcal{HP}_X(C) = \{\, 0.05,\; 0.07,\; 0.09\},$$

reflecting uncertainty across clinical judgment, algorithmic prediction, and laboratory testing.

Example 2.33 (Credit Default Risk from Multiple Agencies)**.** Let Ω be the set of corporate bonds and $(E, \mathcal{E}) = \{\text{Default}, \text{No Default}\}$. Define

$$X(\omega) = \begin{cases} \text{Default}, & \text{if bond } \omega \text{ defaults within one year,} \\ \text{No Default}, & \text{otherwise.} \end{cases}$$

Three rating agencies provide default probabilities:

$$P_{\text{S\&P}}(X = \text{Default}) = 0.02, \quad P_{\text{Moody}}(X = \text{Default}) = 0.03, \quad P_{\text{Fitch}}(X = \text{Default}) = 0.025.$$

Hence for the event $D = \{X = \text{Default}\}$,

$$\mathcal{HP}_X(D) = \{\, 0.02,\; 0.025,\; 0.03\},$$

offering a multi-source view of default risk used in portfolio management.

Received: Received: July 29, 2025 Revised: October 12, 2025 Accepted: December 27, 2025

Theorem 2.34 (Recovery of Classical Distribution)**.** *If* $|K| = 1$*, say* $K = \{k_0\}$*, then* $\mathcal{HP}_X(B) = \{P_{k_0}(X^{-1}(B))\}$ *for all* B*, and the single-valued map* $B \mapsto P_{k_0}(X^{-1}(B))$ *is exactly the ordinary probability distribution of* X*.*

Proof. With $K = \{k_0\}$, Definition 2.31 gives $\mathcal{HP}_X(B) = \{P_{k_0}(X^{-1}(B))\}$. Discarding the singleton bracket yields the classical pushforward measure $X_*P_{k_0}$, which by construction is the law of X under P_{k_0}. Hence hyper-distribution generalizes the usual distribution. $\square$

Theorem 2.35 (Hyper-Additivity)**.** *Let* $\{B_i\}_{i=1}^{\infty} \subseteq \mathcal{E}$ *be pairwise disjoint. Then*

$$\mathcal{HP}_X\Big(\bigcup_{i=1}^{\infty} B_i\Big) = \Big\{\sum_{i=1}^{\infty} P_k\big(X^{-1}(B_i)\big) \;\Big|\; k \in K\Big\}.$$

Moreover, $\mathcal{HP}_X(\varnothing) = \{0\}$ *and* $\mathcal{HP}_X(E) = \{1\}$*.*

Proof. For each $k \in K$, since P_k is a probability measure,

$$P_k\big(X^{-1}(\bigcup_i B_i)\big) = P_k\Big(\bigcup_i X^{-1}(B_i)\Big) = \sum_{i=1}^{\infty} P_k\big(X^{-1}(B_i)\big).$$

Collecting these sums over all k gives the displayed formula. Moreover $X^{-1}(\varnothing) = \varnothing$ and $X^{-1}(E) = \Omega$, so each P_k gives 0 and 1 respectively; hence $\mathcal{HP}_X(\varnothing) = \{0\}$ and $\mathcal{HP}_X(E) = \{1\}$. $\square$

Definition 2.36 ((m, n)-SuperHyperProbability distribution)**.** Let $(\Omega, \mathcal{F})$ and $(E, \mathcal{E})$ be measurable spaces, let

$$\{P_k : \mathcal{F} \to [0, 1]\}_{k \in K}$$

be a nonempty family of probability measures on $(\Omega, \mathcal{F})$, and let

$$X : \Omega \to E$$

be an $\mathcal{F}/\mathcal{E}$-measurable map.

Define the induced hyper-probability distribution

$$\mathrm{HP}_X : \mathcal{E} \longrightarrow \mathcal{P}([0, 1])$$

by

$$\mathrm{HP}_X(B) := \{\, P_k(X^{-1}(B)) \mid k \in K \,\}, \qquad B \in \mathcal{E}.$$

For integers $m \geq 0$ and $n \geq 1$, the (m, n)*-SuperHyperProbability distribution* of X is the map

$$\mathrm{SHPD}_X^{(m,n)} : \mathcal{P}^{(m)}(\mathcal{E}) \longrightarrow \mathcal{P}^{(n)}([0, 1])$$

defined recursively as follows:

(1) **Vertical base and vertical recursion (**$m = 0$**).** For each $B \in \mathcal{E}$,

$$\mathrm{SHPD}_X^{(0,1)}(B) := \mathrm{HP}_X(B),$$

and for $n \geq 2$,

$$\mathrm{SHPD}_X^{(0,n)}(B) := \mathcal{P}\big(\mathrm{SHPD}_X^{(0,n-1)}(B)\big).$$

(2) **Horizontal recursion (**$m \geq 1$**).** For each $U \in \mathcal{P}^{(m)}(\mathcal{E})$, define

$$\mathrm{SHPD}_X^{(m,n)}(U) := \bigcup_{B \in U} \mathrm{SHPD}_X^{(m-1,n)}(B).$$

DOI: https://doi.org/10.54216/PMTCS.060101

Received: Received: July 29, 2025 Revised: October 12, 2025 Accepted: December 27, 2025

Remark 2.37. The ordinary single-valued distribution of X is naturally treated separately as the pushforward measure

$$X_*P(B) := P(X^{-1}(B)), \qquad B \in \mathcal{E}.$$

Thus the present definition starts at the first set-valued level $n = 1$, where $\mathrm{SHPD}_X^{(0,1)} = \mathrm{HP}_X$.

Example 2.38 ($(2,1)$-SuperHyperProbability distribution in weather-event portfolios)**.** Let

$$E = \{\mathrm{Rain}, \mathrm{Storm}\}, \qquad \mathcal{E} = 2^E,$$

and let $X : \Omega \to E$ be a measurable map describing the weather outcome. Set

$$B_R := \{\mathrm{Rain}\}, \qquad B_S := \{\mathrm{Storm}\}.$$

Assume that

$$\mathrm{HP}_X(B_R) = \{0.28, 0.30, 0.32\}, \qquad \mathrm{HP}_X(B_S) = \{0.05, 0.08\}.$$

Define

$$U_1 := \{B_R\}, \qquad U_2 := \{B_S\}, \qquad V := \{U_1, U_2\} \in \mathcal{P}^{(2)}(\mathcal{E}).$$

Then

$$\mathrm{SHPD}_X^{(2,1)}(V) = \mathrm{SHPD}_X^{(1,1)}(U_1) \cup \mathrm{SHPD}_X^{(1,1)}(U_2) = \mathrm{HP}_X(B_R) \cup \mathrm{HP}_X(B_S),$$

so

$$\mathrm{SHPD}_X^{(2,1)}(V) = \{0.05, 0.08, 0.28, 0.30, 0.32\}.$$

This combines all first-level hyper-probability values arising from the two weather-event portfolios.

Example 2.39 ($(1,3)$-SuperHyperProbability distribution for portfolio-loss scenarios)**.** Let

$$E = \{\mathrm{Loss} > 5\%, \mathrm{Loss} \le 5\%\}, \qquad \mathcal{E} = 2^E,$$

and let

$$X : \Omega \to E$$

be the measurable map describing the quarterly loss category. Set

$$C := \{\mathrm{Loss} > 5\%\} \in \mathcal{E}.$$

Assume that four probability measures satisfy

$$P_A(X^{-1}(C)) = 0.10, \quad P_B(X^{-1}(C)) = 0.12, \quad P_C(X^{-1}(C)) = 0.15, \quad P_D(X^{-1}(C)) = 0.08.$$

Then

$$\mathrm{HP}_X(C) = \{0.08, 0.10, 0.12, 0.15\}.$$

Let

$$U := \{C\} \in \mathcal{P}(\mathcal{E}).$$

For $n = 3$,

$$\mathrm{SHPD}_X^{(1,3)}(U) = \mathrm{SHPD}_X^{(0,3)}(C) = \mathcal{P}\Big(\mathcal{P}\big(\{0.08, 0.10, 0.12, 0.15\}\big)\Big).$$

Since $\mathcal{P}(\{0.08, 0.10, 0.12, 0.15\})$ has $2^4 = 16$ elements, the set

$$\mathrm{SHPD}_X^{(1,3)}(U)$$

has 2^{16} elements. For example,

$$\big\{\{0.08, 0.12\}, \{0.10, 0.15\}, \{0.08, 0.10, 0.15\}\big\} \in \mathrm{SHPD}_X^{(1,3)}(U).$$

This is a third-order uncertainty object built from the underlying loss-exceedance assessments.

Theorem 2.40 (Specialization to classical and hyper-probability distributions)**.** *In the setting of Definition 2.36, the following hold.*

Received: Received: July 29, 2025 Revised: October 12, 2025 Accepted: December 27, 2025

(a) *For every* $B \in \mathcal{E}$*,*

$$\mathrm{SHPD}_X^{(0,1)}(B) = \mathrm{HP}_X(B).$$

(b) *If* $K = \{k_0\}$ *is a singleton, then for every* $B \in \mathcal{E}$*,*

$$\mathrm{HP}_X(B) = \{(X_* P_{k_0})(B)\}.$$

Hence $\mathrm{SHPD}_X^{(0,1)}$ *is the singleton-valued embedding of the ordinary probability distribution* $X_* P_{k_0}$*.*

Proof. Part (a) is immediate from the definition:

$$\mathrm{SHPD}_X^{(0,1)}(B) = \mathrm{HP}_X(B) = \{P_k(X^{-1}(B)) \mid k \in K\}.$$

For part (b), if $K = \{k_0\}$, then

$$\mathrm{HP}_X(B) = \{P_{k_0}(X^{-1}(B))\} = \{(X_* P_{k_0})(B)\}.$$

Therefore $\mathrm{SHPD}_X^{(0,1)}$ records exactly the ordinary pushforward probability, viewed as a singleton-valued set. □

Theorem 2.41 (Well-definedness of (m,n)-SuperHyperProbability distributions)**.** *For every* $m \geq 0$ *and* $n \geq 1$*, the map*

$$\mathrm{SHPD}_X^{(m,n)} : \mathcal{P}^{(m)}(\mathcal{E}) \longrightarrow \mathcal{P}^{(n)}([0,1])$$

given in Definition 2.36 is well-defined.

Moreover, for every $n \geq 1$*,*

$$\mathrm{SHPD}_X^{(0,n)}$$

is exactly the usual n*-SuperHyperProbability induced by the hyper-probability distribution* HP_X*.*

Proof. Fix $n \geq 1$. We prove by induction on m that

$$\mathrm{SHPD}_X^{(m,n)}(U) \in \mathcal{P}^{(n)}([0,1]) \qquad \text{for every } U \in \mathcal{P}^{(m)}(\mathcal{E}).$$

Base case: $m = 0$**.** Let $B \in \mathcal{E}$. For $n = 1$,

$$\mathrm{SHPD}_X^{(0,1)}(B) = \mathrm{HP}_X(B) \in \mathcal{P}([0,1]) = \mathcal{P}^{(1)}([0,1]).$$

For $n \geq 2$, assuming inductively that

$$\mathrm{SHPD}_X^{(0,n-1)}(B) \in \mathcal{P}^{(n-1)}([0,1]),$$

we get

$$\mathrm{SHPD}_X^{(0,n)}(B) = \mathcal{P}\big(\mathrm{SHPD}_X^{(0,n-1)}(B)\big) \in \mathcal{P}\big(\mathcal{P}^{(n-1)}([0,1])\big) = \mathcal{P}^{(n)}([0,1]).$$

Inductive step on m**.** Assume that

$$\mathrm{SHPD}_X^{(m,n)} : \mathcal{P}^{(m)}(\mathcal{E}) \to \mathcal{P}^{(n)}([0,1])$$

is well-defined. Let $U \in \mathcal{P}^{(m+1)}(\mathcal{E})$. Then $U \subseteq \mathcal{P}^{(m)}(\mathcal{E})$, so for each $B \in U$,

$$\mathrm{SHPD}_X^{(m,n)}(B) \in \mathcal{P}^{(n)}([0,1]).$$

Since $n \geq 1$, every element of $\mathcal{P}^{(n)}([0,1])$ is a subset of $\mathcal{P}^{(n-1)}([0,1])$. Hence

$$\bigcup_{B \in U} \mathrm{SHPD}_X^{(m,n)}(B) \subseteq \mathcal{P}^{(n-1)}([0,1]),$$

DOI: https://doi.org/10.54216/PMTCS.060101

Received: Received: July 29, 2025 Revised: October 12, 2025 Accepted: December 27, 2025

which implies

$$\bigcup_{B\in U} \mathrm{SHPD}_X^{(m,n)}(B) \in \mathcal{P}\big(\mathcal{P}^{(n-1)}([0,1])\big) = \mathcal{P}^{(n)}([0,1]).$$

Therefore

$$\mathrm{SHPD}_X^{(m+1,n)}(U) = \bigcup_{B\in U} \mathrm{SHPD}_X^{(m,n)}(B)$$

is well-defined.

Thus $\mathrm{SHPD}_X^{(m,n)}$ is well-defined for all $m \geq 0$ and $n \geq 1$. When $m = 0$, the recursive definition is exactly the usual vertical n-SuperHyperProbability construction applied to HP_X. □

Theorem 2.42 (Structural properties of (m,n)-SHPD)**.** *For* $n \geq 1$*, define recursively*

$$Z_1 := \{0\}, \qquad Z_n := \mathcal{P}(Z_{n-1}) \quad (n \geq 2),$$

and

$$O_1 := \{1\}, \qquad O_n := \mathcal{P}(O_{n-1}) \quad (n \geq 2).$$

Also define the m*-fold singleton chain of* E *by*

$$E^{(0)} := E, \qquad E^{(m+1)} := \{E^{(m)}\} \quad (m \geq 0).$$

Then the maps $\mathrm{SHPD}_X^{(m,n)}$ *satisfy the following properties.*

(a) *For* $m = 0$*,*

$$\mathrm{SHPD}_X^{(0,n)}(\varnothing) = Z_n.$$

For every $m \geq 1$*,*

$$\mathrm{SHPD}_X^{(m,n)}(\varnothing) = \varnothing.$$

(b) *For every* $m \geq 0$ *and* $n \geq 1$*,*

$$\mathrm{SHPD}_X^{(m,n)}\big(E^{(m)}\big) = O_n.$$

(c) *(*Additivity at the m-th level*) Let* $m \geq 1$*. For any family* $\{U_j\}_{j\in J}$ *with*

$$U_j \subseteq \mathcal{P}^{(m-1)}(\mathcal{E}) \qquad (j \in J),$$

equivalently $U_j \in \mathcal{P}^{(m)}(\mathcal{E})$*, one has*

$$\mathrm{SHPD}_X^{(m,n)}\Big(\bigcup_{j\in J} U_j\Big) = \bigcup_{j\in J} \mathrm{SHPD}_X^{(m,n)}(U_j).$$

Proof. (a) First let $m = 0$. Since $X^{-1}(\varnothing) = \varnothing$ and every probability measure satisfies $P_k(\varnothing) = 0$, we have

$$\mathrm{SHPD}_X^{(0,1)}(\varnothing) = \mathrm{HP}_X(\varnothing) = \{0\} = Z_1.$$

For $n \geq 2$, by vertical recursion,

$$\mathrm{SHPD}_X^{(0,n)}(\varnothing) = \mathcal{P}\big(\mathrm{SHPD}_X^{(0,n-1)}(\varnothing)\big) = \mathcal{P}(Z_{n-1}) = Z_n.$$

Hence $\mathrm{SHPD}_X^{(0,n)}(\varnothing) = Z_n$ for all $n \geq 1$.

Now let $m \geq 1$. Then

$$\mathrm{SHPD}_X^{(m,n)}(\varnothing) = \bigcup_{B\in\varnothing} \mathrm{SHPD}_X^{(m-1,n)}(B) = \varnothing.$$

(b) For $m \geq 1$, since $E^{(m)} = \{E^{(m-1)}\}$, the horizontal recursion gives

$$\mathrm{SHPD}_X^{(m,n)}\big(E^{(m)}\big) = \bigcup_{B\in E^{(m)}} \mathrm{SHPD}_X^{(m-1,n)}(B) = \mathrm{SHPD}_X^{(m-1,n)}\big(E^{(m-1)}\big).$$

Received: Received: July 29, 2025 Revised: October 12, 2025 Accepted: December 27, 2025

Iterating this identity yields

$$\mathrm{SHPD}_X^{(m,n)}\big(E^{(m)}\big) = \mathrm{SHPD}_X^{(0,n)}(E).$$

Since $X^{-1}(E) = \Omega$ and $P_k(\Omega) = 1$ for all $k \in K$,

$$\mathrm{SHPD}_X^{(0,1)}(E) = \mathrm{HP}_X(E) = \{1\} = O_1.$$

For $n \geq 2$,

$$\mathrm{SHPD}_X^{(0,n)}(E) = \mathcal{P}\big(\mathrm{SHPD}_X^{(0,n-1)}(E)\big) = \mathcal{P}(O_{n-1}) = O_n.$$

Therefore

$$\mathrm{SHPD}_X^{(m,n)}\big(E^{(m)}\big) = O_n$$

for all $m \geq 0$ and $n \geq 1$.

(c) Let $m \geq 1$. Then

$$\mathrm{SHPD}_X^{(m,n)}\Big(\bigcup_{j\in J} U_j\Big) = \bigcup_{B\in\bigcup_{j\in J} U_j} \mathrm{SHPD}_X^{(m-1,n)}(B).$$

Using the elementary identity

$$\bigcup_{B\in\bigcup_{j\in J} U_j} Y_B = \bigcup_{j\in J} \bigcup_{B\in U_j} Y_B,$$

with $Y_B := \mathrm{SHPD}_X^{(m-1,n)}(B)$, we obtain

$$\mathrm{SHPD}_X^{(m,n)}\Big(\bigcup_{j\in J} U_j\Big) = \bigcup_{j\in J} \bigcup_{B\in U_j} \mathrm{SHPD}_X^{(m-1,n)}(B) = \bigcup_{j\in J} \mathrm{SHPD}_X^{(m,n)}(U_j).$$

This proves the claim. □

3 Conclusion

In this paper, we have formally defined the Hyper-Probability Measure and the Hyper-Probability Distribution, examined their fundamental properties, and demonstrated how these constructs unify and extend classical probability theory within the Hyper- and Super-HyperProbability paradigms. Future research will explore potential extensions of these concepts by integrating them with established uncertainty frameworks, including Fuzzy Sets, Intuitionistic Fuzzy Sets, HyperFuzzy Sets, Vague Sets, Rough Sets, Neutrosophic Sets, Hesitant Fuzzy Sets, and Plithogenic Sets. In particular, we expect further progress in the study of extensions based on frameworks such as Neutrosophic Probability.[16–18]

Funding

This study did not receive any financial or external support from organizations or individuals.

Data Availability

This research is purely theoretical, involving no data collection or analysis. We encourage future researchers to pursue empirical investigations to further develop and validate the concepts introduced here.

Ethical Approval

As this research is entirely theoretical in nature and does not involve human participants or animal subjects, no ethical approval is required.

Conflicts of Interest

The authors confirm that there are no conflicts of interest related to the research or its publication.

Use of Generative AI and AI-Assisted Tools

We use generative AI and AI-assisted tools for tasks such as English grammar checking, and We do not employ them in any way that violates ethical standards.

References

[1] Melody Mae Cabigting Lunar and Renson Aguilar Robles. Characterization and structure of a power set graph. *International Journal of Advanced Research and Publications*, 3(6):1–4, 2019.

[2] Florentin Smarandache. Foundation of superhyperstructure & neutrosophic superhyperstructure. *Neutrosophic Sets and Systems*, 63(1):21, 2024.

[3] MA Shalu and S Devi Yamini. Counting maximal independent sets in power set graphs. *Indian Institute of Information Technology Design & Manufacturing (IIITD&M) Kancheepuram, India*, 2014.

[4] Jacob Stegenga. The natural probability theory of stereotypes. *Diametros*, 22(83):26–52, 2025.

[5] Himadri Deshpande. *Foundations of Probability Theory*. Educohack Press, 2025.

[6] Takaaki Fujita. An introduction and reexamination of hyperprobability and superhyperprobability: Comprehensive overview. *Asian Journal of Probability and Statistics*, 27(5):82–109, 2025.

[7] Freddy Delbaen. Coherent risk measures on general probability spaces. 2002.

[8] Jacob Burbea and C. Radhakrishna Rao. Entropy differential metric, distance and divergence measures in probability spaces: A unified approach. *Journal of Multivariate Analysis*, 12:575–596, 1982.

[9] Robert B Ash and Catherine A Doléans-Dade. *Probability and measure theory*. Academic press, 2000.

[10] John S Ramberg, Edward J Dudewicz, Pandu R Tadikamalla, and Edward F Mykytka. A probability distribution and its uses in fitting data. *Technometrics*, 21(2):201–214, 1979.

[11] AJC Wilson. The probability distribution of x-ray intensities. *Acta Crystallographica*, 2(5):318–321, 1949.

[12] David A Nix and Andreas S Weigend. Estimating the mean and variance of the target probability distribution. In *Proceedings of 1994 ieee international conference on neural networks (ICNN'94)*, volume 1, pages 55–60. IEEE, 1994.

[13] Kianoush Nazarpour, Ali H Al-Timemy, Guido Bugmann, and Andrew Jackson. A note on the probability distribution function of the surface electromyogram signal. *Brain research bulletin*, 90:88–91, 2013.

[14] John Gordon Skellam. A probability distribution derived from the binomial distribution by regarding the probability of success as variable between the sets of trials. *Journal of the Royal Statistical Society. Series B (Methodological)*, 10(2):257–261, 1948.

[15] M Edimu, CT Gaunt, and R Herman. Using probability distribution functions in reliability analyses. *Electric Power Systems Research*, 81(4):915–921, 2011.

[16] Jismi Mathew and Milin K Anil. A retrospective review of biomedical applications of neutrosophic probability distributions. *Asian Journal of Probability and Statistics*, 28(2):66–79, 2026.

[17] Hina Khan, Kanwal Javid, et al. A proposed neutrosophic probability model for normalized differencevegetation index using remote sensing: Model building on climate. In *Multi-Criteria Decision Making Models and Techniques: Neutrosophic Approaches*, pages 205–226. IGI Global Scientific Publishing, 2025.

[18] Suman Das, Bimal Shil, Rakhal Das, Huda E Khalid, and AA Salama. Pentapartitioned neutrosophic probability distributions. *Neutrosophic Sets and Systems*, 49:32–47, 2022.

Received: Received: July 29, 2025 Revised: October 12, 2025 Accepted: December 27, 2025