Uncertainty Theory

Third Edition

Baoding Liu

Uncertainty Theory Laboratory

Department of Mathematical Sciences

Tsinghua University

Beijing 100084, China

liu@tsinghua.edu.cn

http://orsc.edu.cn/liu

Lecture Slides at http://orsc.edu.cn/liu/resources.htm

Uncertainty Theory Online at http://orsc.edu.cn/online

Contents

Preface

Most human decisions are made in the state of uncertainty. This fact provides a motivation to study the behavior of uncertain phenomena. In order to model uncertainty, Liu (2007) founded an uncertainty theory that is a branch of mathematics based on normality, monotonicity, self-duality, countable subadditivity, and product measure axioms. Since then, uncertainty theory has been developed steadily and applied widely. Chapter 1 is devoted to the uncertainty theory.

Uncertain programming, proposed by Liu (2009), is a type of mathematical programming involving uncertain variables. Chapter 2 will provide a general framework of uncertain programming, including expected value model, chance-constrained programming, dependent-chance programming, uncertain dynamic programming and uncertain multilevel programming. Chapter 2 will also present some uncertain programming models for system reliability design, project scheduling problem, vehicle routing problem, facility location problem, and machine scheduling problem.

An uncertain process is essentially a sequence of uncertain variables indexed by time or space. Some basic concepts of uncertain process will be presented in Chapter 3.

Uncertain calculus is a branch of mathematics that deals with differentiation and integration of function of uncertain processes. Chapter 4 will introduce the uncertain calculus including canonical process, uncertain integral and chain rule.

Uncertain differential equation is a type of differential equation driven by canonical process. Chapter 5 will discuss the existence, uniqueness and stability of solutions of uncertain differential equations, and present some applications of uncertain differential equation in finance and filtering as well as optimal control.

Uncertain logic is a generalization of mathematical logic for dealing with uncertain knowledge via uncertainty theory. A key point in uncertain logic is that the truth value of an uncertain proposition is defined as the uncertain measure that the proposition is true. One advantage of uncertain logic is the well consistency with classical logic. For example, uncertain logic is consistent with the law of excluded middle and law of contradiction. Chapter 6 deals with uncertain logic.

Uncertain entailment is a methodology for calculating the truth value of an uncertain formula via the maximum uncertainty principle when the truth values of other uncertain formulas are given. Chapter 7 will introduce an

entailment model and discuss modus ponens, modus tollens, and hypothetical syllogism.

Uncertain inference is a process of deriving consequences from uncertain knowledge or evidence via the tool of conditional uncertainty. Chapter 8 will provide some uncertain inference rules, and present applications of uncertain inference in uncertain system and uncertain inference control.

As supplements, Chapter 9 introduces the generalized law of excluded middle and the maximum uncertainty principle. In this chapter, we also provide a brief history of evolution of measures and answer the question "what is uncertainty".

The book is suitable for mathematicians, researchers, engineers, designers, and students in the field of mathematics, information science, operations research, system science, industrial engineering, computer science, artificial intelligence, finance, control, and management science. The readers will learn the axiomatic approach of uncertainty theory, and find this work a stimulating and useful reference.

A Guide for the Reader

The readers are not required to read the book from cover to cover. The logic dependence of chapters is illustrated by the figure below. For exploring probability theory, credibility theory as well as chance theory, the interested reader may continue to read the appendices.
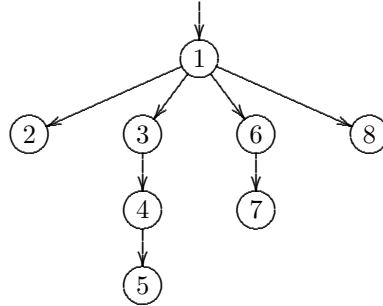
Acknowledgment

Baoding Liu
Tsinghua University
http://orsc.edu.cn/liu

November 4, 2009

Chapter 1

Uncertainty Theory

Most human decisions are made in the state of uncertainty. This fact provides a motivation to study the behavior of uncertain phenomena. In order to model uncertainty, a new mathematical tool is absolutely needed. This is just the *uncertain measure* that is essentially a set function (i.e., a function whose argument is a set) satisfying certain mathematical properties. Uncertain measure is used to measure the truth degree of an uncertain event. In order to develop a theory of uncertain measure, Liu [127] founded an *uncertainty theory* in 2007 that is a branch of mathematics based on normality, monotonicity, self-duality, countable subadditivity, and product measure axioms. Uncertainty theory provides a mathematical model for uncertain phenomena.

The emphasis in this chapter is mainly on uncertain measure, uncertainty space, uncertain variable, identification function, uncertainty distribution, independence, operational law, expected value, variance, moments, critical values, entropy, distance, convergence almost surely, convergence in measure, convergence in mean, convergence in distribution, and conditional uncertainty.

1.1 Uncertainty Space

Let Γ be a nonempty set. A collection $\mathcal{L}$ of subsets of Γ is called a σ-algebra if (a) $\Gamma \in \mathcal{L}$; (b) if $\Lambda \in \mathcal{L}$, then $\Lambda^c \in \mathcal{L}$; and (c) if $\Lambda_1, \Lambda_2, \cdots \in \mathcal{L}$, then $\Lambda_1 \cup \Lambda_2 \cup \cdots \in \mathcal{L}$. Each element Λ in the σ-algebra $\mathcal{L}$ is called an event. Uncertain measure is a function from $\mathcal{L}$ to $[0, 1]$. In order to present an axiomatic definition of uncertain measure, it is necessary to assign to each event Λ a number $\mathcal{M}\{\Lambda\}$ which indicates the truth degree that Λ will occur. In order to ensure that the number $\mathcal{M}\{\Lambda\}$ has certain mathematical properties, Liu [127] proposed the following four axioms:

Axiom 1. *(Normality Axiom)* $\mathcal{M}\{\Gamma\} = 1$ *for the universal set* Γ.

Axiom 2. *(Monotonicity Axiom)* $\mathcal{M}\{\Lambda_1\} \leq \mathcal{M}\{\Lambda_2\}$ *whenever* $\Lambda_1 \subset \Lambda_2$.

Axiom 3. *(Self-Duality Axiom)* $\mathcal{M}\{\Lambda\} + \mathcal{M}\{\Lambda^c\} = 1$ *for any event* Λ.

Axiom 4. *(Countable Subadditivity Axiom) For every countable sequence of events* $\{\Lambda_i\}$*, we have*

$$\mathcal{M}\left\{\bigcup_{i=1}^{\infty} \Lambda_i\right\} \leq \sum_{i=1}^{\infty} \mathcal{M}\{\Lambda_i\}. \tag{1.1}$$

Remark 1.1: The law of contradiction tells us that a proposition cannot be both true and false at the same time, and the law of excluded middle tells us that a proposition is either true or false. Self-duality is in fact a generalization of the law of contradiction and law of excluded middle. In other words, a mathematical system without self-duality assumption will be inconsistent with the laws. This is the main reason why self-duality axiom is assumed.

Remark 1.2: Pathology occurs if subadditivity is not assumed. For example, suppose that a universal set contains 3 elements. We define a set function that takes value 0 for each singleton, and 1 for each set with at least 2 elements. Then such a set function satisfies all axioms but subadditivity. Is it not strange if such a set function serves as a measure?

Remark 1.3: Pathology occurs if countable subadditivity axiom is replaced with finite subadditivity axiom. For example, assume the universal set consists of all real numbers. We define a set function that takes value 0 if the set is bounded, 0.5 if both the set and complement are unbounded, and 1 if the complement of the set is bounded. Then such a set function is finitely subadditive but not countably subadditive. Is it not strange if such a set function serves as a measure? This is the main reason why we accept the countable subadditivity axiom.

Definition 1.1 *(Liu [127]) The set function* $\mathcal{M}$ *is called an uncertain measure if it satisfies the normality, monotonicity, self-duality, and countable subadditivity axioms.*

Example 1.1: Let $\Gamma = \{\gamma_1, \gamma_2, \gamma_3\}$. For this case, there are only 8 events. Define

$$\mathcal{M}\{\gamma_1\} = 0.6, \quad \mathcal{M}\{\gamma_2\} = 0.3, \quad \mathcal{M}\{\gamma_3\} = 0.2,$$

$$\mathcal{M}\{\gamma_1, \gamma_2\} = 0.8, \quad \mathcal{M}\{\gamma_1, \gamma_3\} = 0.7, \quad \mathcal{M}\{\gamma_2, \gamma_3\} = 0.4,$$

$$\mathcal{M}\{\emptyset\} = 0, \quad \mathcal{M}\{\Gamma\} = 1.$$

Then $\mathcal{M}$ is an uncertain measure because it satisfies the four axioms.

Example 1.2: Suppose that $\lambda(x)$ is a nonnegative function on $\Re$ satisfying

$$\sup_{x \neq y} (\lambda(x) + \lambda(y)) = 1. \tag{1.2}$$

Then for any set Λ of real numbers, the set function

$$\mathcal{M}\{\Lambda\} = \begin{cases} \sup_{x \in \Lambda} \lambda(x), & \text{if } \sup_{x \in \Lambda} \lambda(x) < 0.5 \\ 1 - \sup_{x \in \Lambda^c} \lambda(x), & \text{if } \sup_{x \in \Lambda} \lambda(x) \geq 0.5 \end{cases} \tag{1.3}$$

is an uncertain measure on $\Re$.

Example 1.3: Suppose $\rho(x)$ is a nonnegative and integrable function on $\Re$ such that

$$\int_{\Re} \rho(x)\mathrm{d}x \geq 1. \tag{1.4}$$

Then for any Borel set Λ of real numbers, the set function

$$\mathcal{M}\{\Lambda\} = \begin{cases} \int_{\Lambda} \rho(x)\mathrm{d}x, & \text{if } \int_{\Lambda} \rho(x)\mathrm{d}x < 0.5 \\ 1 - \int_{\Lambda^c} \rho(x)\mathrm{d}x, & \text{if } \int_{\Lambda^c} \rho(x)\mathrm{d}x < 0.5 \\ 0.5, & \text{otherwise} \end{cases} \tag{1.5}$$

is an uncertain measure on $\Re$.

Example 1.4: Suppose $\lambda(x)$ is a nonnegative function and $\rho(x)$ is a nonnegative and integrable function on $\Re$ such that

$$\sup_{x \in \Lambda} \lambda(x) + \int_{\Lambda} \rho(x)\mathrm{d}x \geq 0.5 \quad \text{and/or} \quad \sup_{x \in \Lambda^c} \lambda(x) + \int_{\Lambda^c} \rho(x)\mathrm{d}x \geq 0.5 \tag{1.6}$$

for any Borel set Λ of real numbers. Then the set function

$$\mathcal{M}\{\Lambda\} = \begin{cases} \sup_{x \in \Lambda} \lambda(x) + \int_{\Lambda} \rho(x)\mathrm{d}x, & \text{if } \sup_{x \in \Lambda} \lambda(x) + \int_{\Lambda} \rho(x)\mathrm{d}x < 0.5 \\ 1 - \sup_{x \in \Lambda^c} \lambda(x) - \int_{\Lambda^c} \rho(x)\mathrm{d}x, & \text{if } \sup_{x \in \Lambda^c} \lambda(x) + \int_{\Lambda^c} \rho(x)\mathrm{d}x < 0.5 \\ 0.5, & \text{otherwise} \end{cases}$$

is an uncertain measure on $\Re$.

Theorem 1.1 *Suppose that $\mathcal{M}$ is an uncertain measure. Then we have*

$$\mathcal{M}\{\emptyset\} = 0, \tag{1.7}$$

$$0 \leq \mathcal{M}\{\Lambda\} \leq 1 \tag{1.8}$$

for any event Λ.

Proof: It follows from the normality and self-duality axioms that $\mathcal{M}\{\emptyset\} = 1 - \mathcal{M}\{\Gamma\} = 1 - 1 = 0$. It follows from the monotonicity axiom that $0 \leq \mathcal{M}\{\Lambda\} \leq 1$ because $\emptyset \subset \Lambda \subset \Gamma$.

Theorem 1.2 *Suppose that $\mathcal{M}$ is an uncertain measure. Then for any events Λ_1 and Λ_2, we have*

$$\mathcal{M}\{\Lambda_1\} \vee \mathcal{M}\{\Lambda_2\} \leq \mathcal{M}\{\Lambda_1 \cup \Lambda_2\} \leq \mathcal{M}\{\Lambda_1\} + \mathcal{M}\{\Lambda_2\}. \qquad (1.9)$$

Proof: The left-hand inequality follows from the monotonicity axiom and the right-hand inequality follows from the countable subadditivity axiom immediately.

Theorem 1.3 *Suppose that $\mathcal{M}$ is an uncertain measure. Then for any events Λ_1 and Λ_2, we have*

$$\mathcal{M}\{\Lambda_1\} + \mathcal{M}\{\Lambda_2\} - 1 \leq \mathcal{M}\{\Lambda_1 \cap \Lambda_2\} \leq \mathcal{M}\{\Lambda_1\} \wedge \mathcal{M}\{\Lambda_2\}. \qquad (1.10)$$

Proof: The right-hand inequality follows from the monotonicity axiom and the left-hand inequality follows from the self-duality and countable subadditivity axioms, i.e.,

$$\begin{aligned}
\mathcal{M}\{\Lambda_1 \cap \Lambda_2\} &= 1 - \mathcal{M}\{(\Lambda_1 \cap \Lambda_2)^c\} = 1 - \mathcal{M}\{\Lambda_1^c \cup \Lambda_2^c\} \\
&\geq 1 - (\mathcal{M}\{\Lambda_1^c\} + \mathcal{M}\{\Lambda_2^c\}) \\
&= 1 - (1 - \mathcal{M}\{\Lambda_1\}) - (1 - \mathcal{M}\{\Lambda_2\}) \\
&= \mathcal{M}\{\Lambda_1\} + \mathcal{M}\{\Lambda_2\} - 1.
\end{aligned}$$

The inequalities are verified.

Theorem 1.4 *Let $\Gamma = \{\gamma_1, \gamma_2, \cdots\}$. If $\mathcal{M}$ is an uncertain measure, then*

$$\mathcal{M}\{\gamma_i\} + \mathcal{M}\{\gamma_j\} \leq 1 \leq \sum_{k=1}^{\infty} \mathcal{M}\{\gamma_k\} \qquad (1.11)$$

for any i and j.

Proof: Since $\mathcal{M}$ is increasing and self-dual, we have, for any i and j,

$$\mathcal{M}\{\gamma_i\} + \mathcal{M}\{\gamma_j\} \leq \mathcal{M}\{\Gamma \backslash \{\gamma_j\}\} + \mathcal{M}\{\gamma_j\} = 1.$$

Since $\Gamma = \cup_k \{\gamma_k\}$ and $\mathcal{M}$ is countably subadditive, we have

$$1 = \mathcal{M}\{\Gamma\} = \mathcal{M}\left\{ \bigcup_{k=1}^{\infty} \{\gamma_k\} \right\} \leq \sum_{k=1}^{\infty} \mathcal{M}\{\gamma_k\}.$$

The theorem is proved.

Null-Additivity Theorem

Null-additivity is a direct deduction from subadditivity. We first prove a more general theorem.

Theorem 1.5 *Let $\{\Lambda_i\}$ be a sequence of events with $\mathcal{M}\{\Lambda_i\} \to 0$ as $i \to \infty$. Then for any event Λ, we have*

$$\lim_{i\to\infty} \mathcal{M}\{\Lambda \cup \Lambda_i\} = \lim_{i\to\infty} \mathcal{M}\{\Lambda \backslash \Lambda_i\} = \mathcal{M}\{\Lambda\}. \tag{1.12}$$

Proof: It follows from the monotonicity and countable subadditivity axioms that

$$\mathcal{M}\{\Lambda\} \le \mathcal{M}\{\Lambda \cup \Lambda_i\} \le \mathcal{M}\{\Lambda\} + \mathcal{M}\{\Lambda_i\}$$

for each i. Thus we get $\mathcal{M}\{\Lambda \cup \Lambda_i\} \to \mathcal{M}\{\Lambda\}$ by using $\mathcal{M}\{\Lambda_i\} \to 0$. Since $(\Lambda \backslash \Lambda_i) \subset \Lambda \subset ((\Lambda \backslash \Lambda_i) \cup \Lambda_i)$, we have

$$\mathcal{M}\{\Lambda \backslash \Lambda_i\} \le \mathcal{M}\{\Lambda\} \le \mathcal{M}\{\Lambda \backslash \Lambda_i\} + \mathcal{M}\{\Lambda_i\}.$$

Hence $\mathcal{M}\{\Lambda \backslash \Lambda_i\} \to \mathcal{M}\{\Lambda\}$ by using $\mathcal{M}\{\Lambda_i\} \to 0$.

Remark 1.4: It follows from the above theorem that the uncertain measure is null-additive, i.e., $\mathcal{M}\{\Lambda_1 \cup \Lambda_2\} = \mathcal{M}\{\Lambda_1\} + \mathcal{M}\{\Lambda_2\}$ if either $\mathcal{M}\{\Lambda_1\} = 0$ or $\mathcal{M}\{\Lambda_2\} = 0$. In other words, the uncertain measure remains unchanged if the event is enlarged or reduced by an event with uncertain measure zero.

Asymptotic Theorem

Theorem 1.6 *(Asymptotic Theorem) For any events $\Lambda_1, \Lambda_2, \cdots$, we have*

$$\lim_{i\to\infty} \mathcal{M}\{\Lambda_i\} > 0, \quad \textit{if } \Lambda_i \uparrow \Gamma, \tag{1.13}$$

$$\lim_{i\to\infty} \mathcal{M}\{\Lambda_i\} < 1, \quad \textit{if } \Lambda_i \downarrow \emptyset. \tag{1.14}$$

Proof: Assume $\Lambda_i \uparrow \Gamma$. Since $\Gamma = \cup_i \Lambda_i$, it follows from the countable subadditivity axioms that

$$1 = \mathcal{M}\{\Gamma\} \le \sum_{i=1}^{\infty} \mathcal{M}\{\Lambda_i\}.$$

Since $\mathcal{M}\{\Lambda_i\}$ is increasing with respect to i, we have $\lim_{i\to\infty} \mathcal{M}\{\Lambda_i\} > 0$. If $\Lambda_i \downarrow \emptyset$, then $\Lambda_i^c \uparrow \Gamma$. It follows from the first inequality and self-duality axiom that

$$\lim_{i\to\infty} \mathcal{M}\{\Lambda_i\} = 1 - \lim_{i\to\infty} \mathcal{M}\{\Lambda_i^c\} < 1.$$

The theorem is proved.

Example 1.5: Assume Γ is the set of real numbers. Let α be a number with $0 < \alpha \leq 0.5$. Define a set function as follows,

$$\mathcal{M}\{\Lambda\} = \begin{cases} 0, & \text{if } \Lambda = \emptyset \\ \alpha, & \text{if } \Lambda \text{ is upper bounded} \\ 0.5, & \text{if both } \Lambda \text{ and } \Lambda^c \text{ are upper unbounded} \\ 1 - \alpha, & \text{if } \Lambda^c \text{ is upper bounded} \\ 1, & \text{if } \Lambda = \Gamma. \end{cases} \tag{1.15}$$

It is easy to verify that $\mathcal{M}$ is an uncertain measure. Write $\Lambda_i = (-\infty, i]$ for $i = 1, 2, \cdots$ Then $\Lambda_i \uparrow \Gamma$ and $\lim_{i \to \infty} \mathcal{M}\{\Lambda_i\} = \alpha$. Furthermore, we have $\Lambda_i^c \downarrow \emptyset$ and $\lim_{i \to \infty} \mathcal{M}\{\Lambda_i^c\} = 1 - \alpha$.

Uncertainty Space

Definition 1.2 *(Liu [127]) Let Γ be a nonempty set, $\mathcal{L}$ a σ-algebra over Γ, and $\mathcal{M}$ an uncertain measure. Then the triplet $(\Gamma, \mathcal{L}, \mathcal{M})$ is called an uncertainty space.*

Product Measure Axiom and Product Uncertain Measure

Product uncertain measure was defined by Liu [130] in 2009, thus producing the fifth axiom of uncertainty theory called *product measure axiom*. Let $(\Gamma_k, \mathcal{L}_k, \mathcal{M}_k)$ be uncertainty spaces for $k = 1, 2, \cdots, n$. Write

$$\Gamma = \Gamma_1 \times \Gamma_2 \times \cdots \times \Gamma_n, \quad \mathcal{L} = \mathcal{L}_1 \times \mathcal{L}_2 \times \cdots \times \mathcal{L}_n. \tag{1.16}$$

Then there is an uncertain measure $\mathcal{M}$ on the product σ-algebra $\mathcal{L}$ such that

$$\mathcal{M}\{\Lambda_1 \times \Lambda_2 \times \cdots \times \Lambda_n\} = \mathcal{M}_1\{\Lambda_1\} \wedge \mathcal{M}_2\{\Lambda_2\} \wedge \cdots \wedge \mathcal{M}_n\{\Lambda_n\} \tag{1.17}$$

for any measurable rectangle $\Lambda_1 \times \Lambda_2 \times \cdots \times \Lambda_n$. Such an uncertain measure is called the product uncertain measure denoted by

$$\mathcal{M} = \mathcal{M}_1 \wedge \mathcal{M}_2 \wedge \cdots \wedge \mathcal{M}_n. \tag{1.18}$$

In fact, the extension from the class of rectangles to the product σ-algebra $\mathcal{L}$ may be represented as follows.

Axiom 5. *(Liu [130], Product Measure Axiom) Let Γ_k be nonempty sets on which $\mathcal{M}_k$ are uncertain measures, $k = 1, 2, \cdots, n$, respectively. Then the*

product uncertain measure $\mathcal{M} = \mathcal{M}_1 \wedge \mathcal{M}_2 \wedge \cdots \wedge \mathcal{M}_n$ *of each event* $\Lambda \in \mathcal{L}$ *is*

$$
\mathcal{M}\{\Lambda\} =
\begin{cases}
\sup\limits_{\Lambda_1 \times \Lambda_2 \times \cdots \times \Lambda_n \subset \Lambda} \min\limits_{1 \le k \le n} \mathcal{M}_k\{\Lambda_k\}, \\[4pt]
\qquad if \sup\limits_{\Lambda_1 \times \Lambda_2 \times \cdots \times \Lambda_n \subset \Lambda} \min\limits_{1 \le k \le n} \mathcal{M}_k\{\Lambda_k\} > 0.5 \\[4pt]
1 - \sup\limits_{\Lambda_1 \times \Lambda_2 \times \cdots \times \Lambda_n \subset \Lambda^c} \min\limits_{1 \le k \le n} \mathcal{M}_k\{\Lambda_k\}, \\[4pt]
\qquad if \sup\limits_{\Lambda_1 \times \Lambda_2 \times \cdots \times \Lambda_n \subset \Lambda^c} \min\limits_{1 \le k \le n} \mathcal{M}_k\{\Lambda_k\} > 0.5 \\[4pt]
0.5, \qquad otherwise.
\end{cases}
\tag{1.19}
$$

Figure 1.1: Graphical Illustration of Extension from the Class of Rectangles to the Product σ-Algebra. The uncertain measure of Λ (the disk) is essentially the acreage of its inscribed rectangle $\Lambda_1 \times \Lambda_2$ if it is greater than 0.5. Otherwise, we have to examine its complement Λ^c. If the inscribed rectangle of Λ^c is greater than 0.5, then $\mathcal{M}\{\Lambda^c\}$ is just its inscribed rectangle and $\mathcal{M}\{\Lambda\} = 1 - \mathcal{M}\{\Lambda^c\}$. If there does not exist an inscribed rectangle of Λ or Λ^c greater than 0.5, then we set $\mathcal{M}\{\Lambda\} = 0.5$.

Theorem 1.7 *(Peng [185]) The product uncertain measure (1.19) is an uncertain measure.*

Proof: In order to prove that the product uncertain measure (1.19) is indeed an uncertain measure, we should verify that the product uncertain measure satisfies the normality, monotonicity, self-duality and countable subadditivity axioms.

STEP 1: At first, for any event $\Lambda \in \mathcal{L}$, it is easy to verify that

$$
\sup\limits_{\Lambda_1 \times \Lambda_2 \times \cdots \times \Lambda_n \subset \Lambda} \min\limits_{1 \le k \le n} \mathcal{M}_k\{\Lambda_k\} + \sup\limits_{\Lambda_1 \times \Lambda_2 \times \cdots \times \Lambda_n \subset \Lambda^c} \min\limits_{1 \le k \le n} \mathcal{M}_k\{\Lambda_k\} \le 1.
$$

This means that at most one of

$$\sup_{\Lambda_1 \times \Lambda_2 \times \cdots \times \Lambda_n \subset \Lambda} \min_{1 \leq k \leq n} \mathcal{M}_k\{\Lambda_k\} \quad \text{and} \quad \sup_{\Lambda_1 \times \Lambda_2 \times \cdots \times \Lambda_n \subset \Lambda^c} \min_{1 \leq k \leq n} \mathcal{M}_k\{\Lambda_k\}$$

is greater than 0.5. Thus the expression (1.19) is reasonable.

STEP 2: The product uncertain measure is clearly normal, i.e., $\mathcal{M}\{\Gamma\} = 1$.

STEP 3: We prove the self-duality, i.e., $\mathcal{M}\{\Lambda\} + \mathcal{M}\{\Lambda^c\} = 1$. The argument breaks down into three cases. Case 1: Assume

$$\sup_{\Lambda_1 \times \Lambda_2 \times \cdots \times \Lambda_n \subset \Lambda} \min_{1 \leq k \leq n} \mathcal{M}_k\{\Lambda_k\} > 0.5.$$

Then we immediately have

$$\sup_{\Lambda_1 \times \Lambda_2 \times \cdots \times \Lambda_n \subset \Lambda^c} \min_{1 \leq k \leq n} \mathcal{M}_k\{\Lambda_k\} < 0.5.$$

It follows from (1.19) that

$$\mathcal{M}\{\Lambda\} = \sup_{\Lambda_1 \times \Lambda_2 \times \cdots \times \Lambda_n \subset \Lambda} \min_{1 \leq k \leq n} \mathcal{M}_k\{\Lambda_k\},$$

$$\mathcal{M}\{\Lambda^c\} = 1 - \sup_{\Lambda_1 \times \Lambda_2 \times \cdots \times \Lambda_n \subset (\Lambda^c)^c} \min_{1 \leq k \leq n} \mathcal{M}_k\{\Lambda_k\} = 1 - \mathcal{M}\{\Lambda\}.$$

The self-duality is proved. Case 2: Assume

$$\sup_{\Lambda_1 \times \Lambda_2 \times \cdots \times \Lambda_n \subset \Lambda^c} \min_{1 \leq k \leq n} \mathcal{M}_k\{\Lambda_k\} > 0.5.$$

This case may be proved by a similar process. Case 3: Assume

$$\sup_{\Lambda_1 \times \Lambda_2 \times \cdots \times \Lambda_n \subset \Lambda} \min_{1 \leq k \leq n} \mathcal{M}_k\{\Lambda_k\} \leq 0.5$$

and

$$\sup_{\Lambda_1 \times \Lambda_2 \times \cdots \times \Lambda_n \subset \Lambda^c} \min_{1 \leq k \leq n} \mathcal{M}_k\{\Lambda_k\} \leq 0.5.$$

It follows from (1.19) that $\mathcal{M}\{\Lambda\} = \mathcal{M}\{\Lambda^c\} = 0.5$ which proves the self-duality.

STEP 4: Let us prove that $\mathcal{M}$ is increasing. Suppose Λ and Δ are two events in $\mathcal{L}$ with $\Lambda \subset \Delta$. The argument breaks down into three cases. Case 1: Assume

$$\sup_{\Lambda_1 \times \Lambda_2 \times \cdots \times \Lambda_n \subset \Lambda} \min_{1 \leq k \leq n} \mathcal{M}_k\{\Lambda_k\} > 0.5.$$

Then

$$\sup_{\Delta_1 \times \Delta_2 \times \cdots \times \Delta_n \subset \Delta} \min_{1 \leq k \leq n} \mathcal{M}_k\{\Delta_k\} \geq \sup_{\Lambda_1 \times \Lambda_2 \times \cdots \times \Lambda_n \subset \Lambda} \min_{1 \leq k \leq n} \mathcal{M}_k\{\Lambda_k\} > 0.5.$$

It follows from (1.19) that $\mathcal{M}\{\Lambda\} \leq \mathcal{M}\{\Delta\}$. Case 2: Assume

$$\sup_{\Delta_1 \times \Delta_2 \times \cdots \times \Delta_n \subset \Delta^c} \min_{1 \leq k \leq n} \mathcal{M}_k\{\Delta_k\} > 0.5.$$

Then

$$\sup_{\Lambda_1 \times \Lambda_2 \times \cdots \times \Lambda_n \subset \Lambda^c} \min_{1 \leq k \leq n} \mathcal{M}_k\{\Lambda_k\} \geq \sup_{\Delta_1 \times \Delta_2 \times \cdots \times \Delta_n \subset \Delta^c} \min_{1 \leq k \leq n} \mathcal{M}_k\{\Delta_k\} > 0.5.$$

Thus

$$\mathcal{M}\{\Lambda\} = 1 - \sup_{\Lambda_1 \times \Lambda_2 \times \cdots \times \Lambda_n \subset \Lambda^c} \min_{1 \leq k \leq n} \mathcal{M}_k\{\Lambda_k\}$$

$$\leq 1 - \sup_{\Delta_1 \times \Delta_2 \times \cdots \times \Delta_n \subset \Delta^c} \min_{1 \leq k \leq n} \mathcal{M}_k\{\Delta_k\} = \mathcal{M}\{\Delta\}.$$

Case 3: Assume

$$\sup_{\Lambda_1 \times \Lambda_2 \times \cdots \times \Lambda_n \subset \Lambda} \min_{1 \leq k \leq n} \mathcal{M}_k\{\Lambda_k\} \leq 0.5$$

and

$$\sup_{\Delta_1 \times \Delta_2 \times \cdots \times \Delta_n \subset \Delta^c} \min_{1 \leq k \leq n} \mathcal{M}_k\{\Delta_k\} \leq 0.5.$$

Then

$$\mathcal{M}\{\Lambda\} \leq 0.5 \leq 1 - \mathcal{M}\{\Delta^c\} = \mathcal{M}\{\Delta\}.$$

STEP 5: Finally, we prove the countable subadditivity of $\mathcal{M}$. For simplicity, we only prove the case of two events Λ and Δ. The argument breaks down into three cases. Case 1: Assume $\mathcal{M}\{\Lambda\} < 0.5$ and $\mathcal{M}\{\Delta\} < 0.5$. For any given $\varepsilon > 0$, there are two rectangles

$$\Lambda_1 \times \Lambda_2 \times \cdots \times \Lambda_n \subset \Lambda^c, \quad \Delta_1 \times \Delta_2 \times \cdots \times \Delta_n \subset \Delta^c$$

such that
$$1 - \min_{1 \leq k \leq n} \mathcal{M}_k\{\Lambda_k\} \leq \mathcal{M}\{\Lambda\} + \varepsilon/2,$$

$$1 - \min_{1 \leq k \leq n} \mathcal{M}_k\{\Delta_k\} \leq \mathcal{M}\{\Delta\} + \varepsilon/2.$$

Note that

$$(\Lambda_1 \cap \Delta_1) \times (\Lambda_2 \cap \Delta_2) \times \cdots \times (\Lambda_n \cap \Delta_n) \subset (\Lambda \cup \Delta)^c.$$

It follows from Theorem 1.3 that

$$\mathcal{M}_k\{\Lambda_k \cap \Delta_k\} \geq \mathcal{M}_k\{\Lambda_k\} + \mathcal{M}_k\{\Delta_k\} - 1$$

for any k. Thus

$$\mathcal{M}\{\Lambda \cup \Delta\} \leq 1 - \min_{1 \leq k \leq n} \mathcal{M}_k\{\Lambda_k \cap \Delta_k\}$$

$$\leq 1 - \min_{1 \leq k \leq n} \mathcal{M}_k\{\Lambda_k\} + 1 - \min_{1 \leq k \leq n} \mathcal{M}_k\{\Delta_k\}$$

$$\leq \mathcal{M}\{\Lambda\} + \mathcal{M}\{\Delta\} + \varepsilon.$$

Letting $\varepsilon \to 0$, we obtain

$$\mathcal{M}\{\Lambda \cup \Delta\} \leq \mathcal{M}\{\Lambda\} + \mathcal{M}\{\Delta\}.$$

Case 2: Assume $\mathcal{M}\{\Lambda\} \geq 0.5$ and $\mathcal{M}\{\Delta\} < 0.5$. When $\mathcal{M}\{\Lambda \cup \Delta\} = 0.5$, the subadditivity is obvious. Now we consider the case $\mathcal{M}\{\Lambda \cup \Delta\} > 0.5$, i.e., $\mathcal{M}\{\Lambda^c \cap \Delta^c\} < 0.5$. By using $\Lambda^c \cup \Delta = (\Lambda^c \cap \Delta^c) \cup \Delta$ and Case 1, we get

$$\mathcal{M}\{\Lambda^c \cup \Delta\} \leq \mathcal{M}\{\Lambda^c \cap \Delta^c\} + \mathcal{M}\{\Delta\}.$$

Thus

$$\mathcal{M}\{\Lambda \cup \Delta\} = 1 - \mathcal{M}\{\Lambda^c \cap \Delta^c\} \leq 1 - \mathcal{M}\{\Lambda^c \cup \Delta\} + \mathcal{M}\{\Delta\}$$

$$\leq 1 - \mathcal{M}\{\Lambda^c\} + \mathcal{M}\{\Delta\} = \mathcal{M}\{\Lambda\} + \mathcal{M}\{\Delta\}.$$

Case 3: If both $\mathcal{M}\{\Lambda\} \geq 0.5$ and $\mathcal{M}\{\Delta\} \geq 0.5$, then the subadditivity is obvious because $\mathcal{M}\{\Lambda\} + \mathcal{M}\{\Delta\} \geq 1$. The theorem is proved.

Definition 1.3 *Let* $(\Gamma_k, \mathcal{L}_k, \mathcal{M}_k), k = 1, 2, \cdots, n$ *be uncertainty spaces,* $\Gamma = \Gamma_1 \times \Gamma_2 \times \cdots \times \Gamma_n$, $\mathcal{L} = \mathcal{L}_1 \times \mathcal{L}_2 \times \cdots \times \mathcal{L}_n$, *and* $\mathcal{M} = \mathcal{M}_1 \wedge \mathcal{M}_2 \wedge \cdots \wedge \mathcal{M}_n$. *Then* $(\Gamma, \mathcal{L}, \mathcal{M})$ *is called the product uncertainty space of* $(\Gamma_k, \mathcal{L}_k, \mathcal{M}_k), k = 1, 2, \cdots, n$.

1.2   Uncertain Variable

Definition 1.4 *(Liu [127]) An uncertain variable is a measurable function* ξ *from an uncertainty space* $(\Gamma, \mathcal{L}, \mathcal{M})$ *to the set of real numbers, i.e., for any Borel set* B *of real numbers, the set*

$$\{\xi \in B\} = \{\gamma \in \Gamma \mid \xi(\gamma) \in B\} \tag{1.20}$$

is an event.

Example 1.6: Take $(\Gamma, \mathcal{L}, \mathcal{M})$ to be $\{\gamma_1, \gamma_2\}$ with $\mathcal{M}\{\gamma_1\} = \mathcal{M}\{\gamma_2\} = 0.5$. Then the function

$$\xi(\gamma) = \begin{cases} 0, & \text{if } \gamma = \gamma_1 \\ 1, & \text{if } \gamma = \gamma_2 \end{cases}$$

is an uncertain variable.

Example 1.7: A crisp number c may be regarded as a special uncertain variable. In fact, it is the constant function $\xi(\gamma) \equiv c$ on the uncertainty space $(\Gamma, \mathcal{L}, \mathcal{M})$.
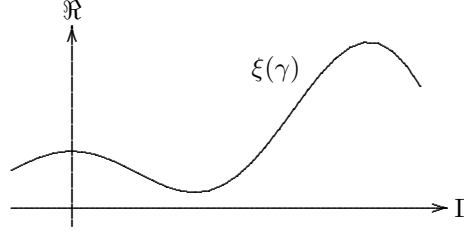
Figure 1.2: An Uncertain Variable

Definition 1.5 *An uncertain variable ξ on the uncertainty space $(\Gamma, \mathcal{L}, \mathcal{M})$ is said to be*
(a) nonnegative if $\mathcal{M}\{\xi < 0\} = 0$;
(b) positive if $\mathcal{M}\{\xi \leq 0\} = 0$;
(c) continuous if $\mathcal{M}\{\xi = x\}$ is a continuous function of x;
(d) simple if there exists a finite sequence $\{x_1, x_2, \cdots, x_m\}$ such that

$$\mathcal{M}\{\xi \neq x_1, \xi \neq x_2, \cdots, \xi \neq x_m\} = 0; \tag{1.21}$$

(e) discrete if there exists a countable sequence $\{x_1, x_2, \cdots\}$ such that

$$\mathcal{M}\{\xi \neq x_1, \xi \neq x_2, \cdots\} = 0. \tag{1.22}$$

Definition 1.6 *Let ξ and η be uncertain variables defined on the uncertainty space $(\Gamma, \mathcal{L}, \mathcal{M})$. We say $\xi = \eta$ if $\xi(\gamma) = \eta(\gamma)$ for almost all $\gamma \in \Gamma$.*

Definition 1.7 *The uncertain variables ξ and η are identically distributed if*

$$\mathcal{M}\{\xi \in B\} = \mathcal{M}\{\eta \in B\} \tag{1.23}$$

for any Borel set B of real numbers.

It is clear that uncertain variables ξ and η are identically distributed if $\xi = \eta$. However, identical distribution does not imply $\xi = \eta$. For example, let $(\Gamma, \mathcal{L}, \mathcal{M})$ be $\{\gamma_1, \gamma_2\}$ with $\mathcal{M}\{\gamma_1\} = \mathcal{M}\{\gamma_2\} = 0.5$. Define

$$\xi(\gamma) = \begin{cases} 1, & \text{if } \gamma = \gamma_1 \\ -1, & \text{if } \gamma = \gamma_2, \end{cases} \qquad \eta(\gamma) = \begin{cases} -1, & \text{if } \gamma = \gamma_1 \\ 1, & \text{if } \gamma = \gamma_2. \end{cases}$$

The two uncertain variables ξ and η are identically distributed but $\xi \neq \eta$.

Uncertain Vector

Definition 1.8 *An n-dimensional uncertain vector is a measurable function from an uncertainty space $(\Gamma, \mathcal{L}, \mathcal{M})$ to the set of n-dimensional real vectors, i.e., for any Borel set B of $\Re^n$, the set*

$$\{\boldsymbol{\xi} \in B\} = \left\{ \gamma \in \Gamma \mid \boldsymbol{\xi}(\gamma) \in B \right\} \tag{1.24}$$

is an event.

Theorem 1.8 *The vector* $(\xi_1, \xi_2, \cdots, \xi_n)$ *is an uncertain vector if and only if* $\xi_1, \xi_2, \cdots, \xi_n$ *are uncertain variables.*

Proof: Write $\boldsymbol{\xi} = (\xi_1, \xi_2, \cdots, \xi_n)$. Suppose that $\boldsymbol{\xi}$ is an uncertain vector on the uncertainty space $(\Gamma, \mathcal{L}, \mathcal{M})$. For any Borel set B of $\Re$, the set $B \times \Re^{n-1}$ is a Borel set of $\Re^n$. Thus the set

$$\{\xi_1 \in B\} = \{\xi_1 \in B, \xi_2 \in \Re, \cdots, \xi_n \in \Re\} = \{\boldsymbol{\xi} \in B \times \Re^{n-1}\}$$

is an event. Hence ξ_1 is an uncertain variable. A similar process may prove that $\xi_2, \xi_3, \cdots, \xi_n$ are uncertain variables. Conversely, suppose that all $\xi_1, \xi_2, \cdots, \xi_n$ are uncertain variables on the uncertainty space $(\Gamma, \mathcal{L}, \mathcal{M})$. We define

$$\mathcal{B} = \left\{ B \subset \Re^n \mid \{\boldsymbol{\xi} \in B\} \text{ is an event} \right\}.$$

The vector $\boldsymbol{\xi} = (\xi_1, \xi_2, \cdots, \xi_n)$ is proved to be an uncertain vector if we can prove that $\mathcal{B}$ contains all Borel sets of $\Re^n$. First, the class $\mathcal{B}$ contains all open intervals of $\Re^n$ because

$$\left\{ \boldsymbol{\xi} \in \prod_{i=1}^{n}(a_i, b_i) \right\} = \bigcap_{i=1}^{n} \{\xi_i \in (a_i, b_i)\}$$

is an event. Next, the class $\mathcal{B}$ is a σ-algebra of $\Re^n$ because (i) we have $\Re^n \in \mathcal{B}$ since $\{\boldsymbol{\xi} \in \Re^n\} = \Gamma$; (ii) if $B \in \mathcal{B}$, then $\{\boldsymbol{\xi} \in B\}$ is an event, and

$$\{\boldsymbol{\xi} \in B^c\} = \{\boldsymbol{\xi} \in B\}^c$$

is an event. This means that $B^c \in \mathcal{B}$; (iii) if $B_i \in \mathcal{B}$ for $i = 1, 2, \cdots$, then $\{\boldsymbol{\xi} \in B_i\}$ are events and

$$\left\{ \boldsymbol{\xi} \in \bigcup_{i=1}^{\infty} B_i \right\} = \bigcup_{i=1}^{\infty} \{\boldsymbol{\xi} \in B_i\}$$

is an event. This means that $\cup_i B_i \in \mathcal{B}$. Since the smallest σ-algebra containing all open intervals of $\Re^n$ is just the Borel algebra of $\Re^n$, the class $\mathcal{B}$ contains all Borel sets of $\Re^n$. The theorem is proved.

Uncertain Arithmetic

Definition 1.9 *Suppose that* $f : \Re^n \to \Re$ *is a measurable function, and* $\xi_1, \xi_2, \cdots, \xi_n$ *uncertain variables on the uncertainty space* $(\Gamma, \mathcal{L}, \mathcal{M})$. *Then* $\xi = f(\xi_1, \xi_2, \cdots, \xi_n)$ *is an uncertain variable defined as*

$$\xi(\gamma) = f(\xi_1(\gamma), \xi_2(\gamma), \cdots, \xi_n(\gamma)), \quad \forall \gamma \in \Gamma. \tag{1.25}$$

Example 1.8: Let ξ_1 and ξ_2 be two uncertain variables. Then the sum $\xi = \xi_1 + \xi_2$ is an uncertain variable defined by

$$\xi(\gamma) = \xi_1(\gamma) + \xi_2(\gamma), \quad \forall \gamma \in \Gamma.$$

The product $\xi = \xi_1 \xi_2$ is also an uncertain variable defined by

$$\xi(\gamma) = \xi_1(\gamma) \cdot \xi_2(\gamma), \quad \forall \gamma \in \Gamma.$$

The reader may wonder whether $\xi(\gamma_1, \gamma_2, \cdots, \gamma_n)$ defined by (1.25) is an uncertain variable. The following theorem answers this question.

Theorem 1.9 *Let $\boldsymbol{\xi}$ be an n-dimensional uncertain vector, and $f : \Re^n \to \Re$ a measurable function. Then $f(\boldsymbol{\xi})$ is an uncertain variable such that*

$$\mathcal{M}\{f(\boldsymbol{\xi}) \in B\} = \mathcal{M}\{\boldsymbol{\xi} \in f^{-1}(B)\} \tag{1.26}$$

for any Borel set B of real numbers.

Proof: Assume that $\boldsymbol{\xi}$ is an uncertain vector on the uncertainty space $(\Gamma, \mathcal{L}, \mathcal{M})$. For any Borel set B of $\Re$, since f is a measurable function, the $f^{-1}(B)$ is a Borel set of $\Re^n$. Thus the set $\{f(\boldsymbol{\xi}) \in B\} = \{\boldsymbol{\xi} \in f^{-1}(B)\}$ is an event for any Borel set B. Hence $f(\boldsymbol{\xi})$ is an uncertain variable.

1.3   Identification Functions

This section will introduce three types of *identification functions* to characterize uncertain variables.

First Identification Function

Definition 1.10 *An uncertain variable ξ is said to have a first identification function λ if*
(i) $\lambda(x)$ is a nonnegative function on $\Re$ such that

$$\sup_{x \neq y} (\lambda(x) + \lambda(y)) = 1; \tag{1.27}$$

(ii) for any set B of real numbers, we have

$$\mathcal{M}\{\xi \in B\} = \begin{cases} \sup_{x \in B} \lambda(x), & \text{if } \sup_{x \in B} \lambda(x) < 0.5 \\[2mm] 1 - \sup_{x \in B^c} \lambda(x), & \text{if } \sup_{x \in B} \lambda(x) \geq 0.5. \end{cases} \tag{1.28}$$

Example 1.9: By a *rectangular uncertain variable* we mean the uncertain variable fully determined by the pair (a, b) of crisp numbers with $a < b$, whose first identification function is

$$\lambda_1(x) = 0.5, \quad a \leq x \leq b.$$

Example 1.10: By a *triangular uncertain variable* we mean the uncertain variable fully determined by the triplet (a, b, c) of crisp numbers with $a < b < c$, whose first identification function is

$$\lambda_2(x) = \begin{cases} \dfrac{x-a}{2(b-a)}, & \text{if } a \leq x \leq b \\[2mm] \dfrac{x-c}{2(b-c)}, & \text{if } b \leq x \leq c. \end{cases}$$

Example 1.11: By a *trapezoidal uncertain variable* we mean the uncertain variable fully determined by the quadruplet (a, b, c, d) of crisp numbers with $a < b < c < d$, whose first identification function is

$$\lambda_3(x) = \begin{cases} \dfrac{x-a}{2(b-a)}, & \text{if } a \leq x \leq b \\[2mm] 0.5, & \text{if } b \leq x \leq c \\[2mm] \dfrac{x-d}{2(c-d)}, & \text{if } c \leq x \leq d. \end{cases}$$
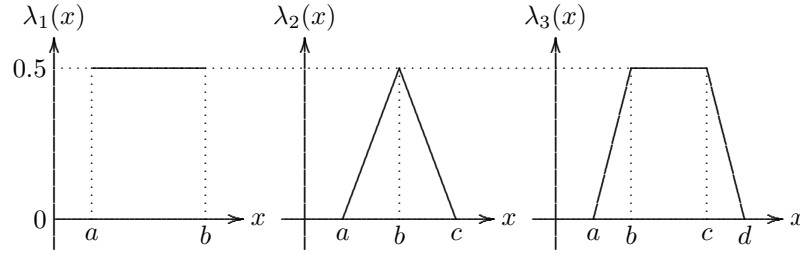
Figure 1.3: First Identification Functions λ_1, λ_2 and λ_3

Theorem 1.10 *Suppose λ is a nonnegative function satisfying (1.27). Then there is an uncertain variable ξ such that (1.28) holds.*

Proof: Let $\Re$ be the universal set. For each set B of real numbers, we define a set function

$$\mathcal{M}\{B\} = \begin{cases} \sup_{x \in B} \lambda(x), & \text{if } \sup_{x \in B} \lambda(x) < 0.5 \\[3mm] 1 - \sup_{x \in B^c} \lambda(x), & \text{if } \sup_{x \in B} \lambda(x) \geq 0.5. \end{cases}$$

It is clear that $\mathcal{M}$ is normal, increasing, self-dual, and countably subadditive. That is, the set function $\mathcal{M}$ is indeed an uncertain measure. Now we define an uncertain variable ξ as an identity function from the uncertainty space $(\Re, \mathcal{A}, \mathcal{M})$ to $\Re$. We may verify that ξ meets (1.28). The theorem is proved.

Theorem 1.11 *(First Measure Inversion Theorem) Let ξ be an uncertain variable with first identification function λ. Then for any set B of real numbers, we have*

$$\mathcal{M}\{\xi \in B\} = \begin{cases} \sup_{x \in B} \lambda(x), & if \; \sup_{x \in B} \lambda(x) < 0.5 \\ 1 - \sup_{x \in B^c} \lambda(x), & if \; \sup_{x \in B} \lambda(x) \geq 0.5. \end{cases} \tag{1.29}$$

Proof: It follows from the definition of first identification function immediately.

Example 1.12: Suppose ξ is an uncertain variable with first identification function λ. Then

$$\mathcal{M}\{\xi = x\} = \lambda(x), \quad \forall x \in \Re.$$

Example 1.13: Suppose ξ is a discrete uncertain variable taking values in $\{x_1, x_2, \cdots\}$ with *first identification coefficients* $\{\lambda_1, \lambda_2, \cdots\}$. Note that the coefficients must meet

$$\sup_{i \neq j} (\lambda_i + \lambda_j) = 1. \tag{1.30}$$

For any set B of real numbers, we have

$$\mathcal{M}\{\xi \in B\} = \begin{cases} \sup_{x_i \in B} \lambda_i, & if \; \sup_{x_i \in B} \lambda_i < 0.5 \\ 1 - \sup_{x_i \in B^c} \lambda_i, & if \; \sup_{x_i \in B} \lambda_i \geq 0.5. \end{cases} \tag{1.31}$$

Second Identification Function

Definition 1.11 *(Liu [137]) An uncertain variable ξ is said to have a second identification function ρ if*
(i) $\rho(x)$ is a nonnegative and integrable function on $\Re$ such that

$$\int_{\Re} \rho(x)\mathrm{d}x \geq 1; \tag{1.32}$$

(ii) for any Borel set B of real numbers, we have

$$\mathcal{M}\{\xi \in B\} = \begin{cases} \displaystyle\int_B \rho(x)\mathrm{d}x, & if \; \displaystyle\int_B \rho(x)\mathrm{d}x < 0.5 \\ 1 - \displaystyle\int_{B^c} \rho(x)\mathrm{d}x, & if \; \displaystyle\int_{B^c} \rho(x)\mathrm{d}x < 0.5 \\ 0.5, & otherwise. \end{cases} \tag{1.33}$$

Theorem 1.12 *Suppose ρ is a nonnegative and integrable function satisfying (1.32). Then there is an uncertain variable ξ such that (1.33) holds.*

Proof: Let $\Re$ be the universal set. For each Borel set B of real numbers, we define a set function

$$\mathcal{M}\{B\} = \begin{cases} \displaystyle\int_B \rho(x)\mathrm{d}x, & \text{if } \displaystyle\int_B \rho(x)\mathrm{d}x < 0.5 \\[2ex] 1 - \displaystyle\int_{B^c} \rho(x)\mathrm{d}x, & \text{if } \displaystyle\int_{B^c} \rho(x)\mathrm{d}x < 0.5 \\[2ex] 0.5, & \text{otherwise.} \end{cases}$$

It is clear that $\mathcal{M}$ is normal, increasing, self-dual, and countably subadditive. That is, the set function $\mathcal{M}$ is indeed an uncertain measure. Now we define an uncertain variable ξ as an identity function from the uncertainty space $(\Re, \mathcal{A}, \mathcal{M})$ to $\Re$. We may verify that ξ meets (1.33). The theorem is proved.

Theorem 1.13 *(Second Measure Inversion Theorem) Let ξ be an uncertain variable with second identification function ρ. Then for any Borel set B of real numbers, we have*

$$\mathcal{M}\{\xi \in B\} = \begin{cases} \displaystyle\int_B \rho(x)\mathrm{d}x, & \text{if } \displaystyle\int_B \rho(x)\mathrm{d}x < 0.5 \\[2ex] 1 - \displaystyle\int_{B^c} \rho(x)\mathrm{d}x, & \text{if } \displaystyle\int_{B^c} \rho(x)\mathrm{d}x < 0.5 \\[2ex] 0.5, & \text{otherwise.} \end{cases} \tag{1.34}$$

Proof: It follows from the definition of second identification function immediately.

Example 1.14: Suppose ξ is a discrete uncertain variable taking values in $\{x_1, x_2, \cdots\}$ with *second identification coefficients* $\{\rho_1, \rho_2, \cdots\}$. Note that the coefficients must meet

$$\sum_{i=1}^{\infty} \rho_i \geq 1. \tag{1.35}$$

For any set B of real numbers, we have

$$\mathcal{M}\{\xi \in B\} = \begin{cases} \displaystyle\sum_{x_i \in B} \rho_i, & \text{if } \displaystyle\sum_{x_i \in B} \rho_i < 0.5 \\[2ex] 1 - \displaystyle\sum_{x_i \in B^c} \rho_i, & \text{if } \displaystyle\sum_{x_i \in B^c} \rho_i < 0.5 \\[2ex] 0.5, & \text{otherwise.} \end{cases} \tag{1.36}$$

Third Identification Function

Definition 1.12 *(Liu [130]) An uncertain variable ξ is said to have a third identification function (λ, ρ) if (i) $\lambda(x)$ is a nonnegative function and $\rho(x)$ is*

a nonnegative and integrable function such that

$$\sup_{x \in B} \lambda(x) + \int_B \rho(x)\mathrm{d}x \ge 0.5 \quad and/or \quad \sup_{x \in B^c} \lambda(x) + \int_{B^c} \rho(x)\mathrm{d}x \ge 0.5 \quad (1.37)$$

for any Borel set B of real numbers; (ii) we have

$$\mathcal{M}\{\xi \in B\} = \begin{cases} \sup_{x \in B} \lambda(x) + \int_B \rho(x)\mathrm{d}x, \\ \qquad if \ \sup_{x \in B} \lambda(x) + \int_B \rho(x)\mathrm{d}x < 0.5 \\ 1 - \sup_{x \in B^c} \lambda(x) - \int_{B^c} \rho(x)\mathrm{d}x, \\ \qquad if \ \sup_{x \in B^c} \lambda(x) + \int_{B^c} \rho(x)\mathrm{d}x < 0.5 \\ 0.5, \qquad otherwise. \end{cases} \quad (1.38)$$

Theorem 1.14 *Suppose λ is a nonnegative function and ρ is a nonnegative and integrable function satisfying the condition (1.37). Then there is an uncertain variable ξ such that (1.38) holds.*

Proof: Let $\Re$ be the universal set. For each Borel set B of real numbers, we define a set function

$$\mathcal{M}\{B\} = \begin{cases} \sup_{x \in B} \lambda(x) + \int_B \rho(x)\mathrm{d}x, & if \ \sup_{x \in B} \lambda(x) + \int_B \rho(x)\mathrm{d}x < 0.5 \\ 1 - \sup_{x \in B^c} \lambda(x) - \int_{B^c} \rho(x)\mathrm{d}x, & if \ \sup_{x \in B^c} \lambda(x) + \int_{B^c} \rho(x)\mathrm{d}x < 0.5 \\ 0.5, & otherwise. \end{cases}$$

It is clear that $\mathcal{M}$ is normal, increasing, self-dual, and countably subadditive. That is, the set function $\mathcal{M}$ is indeed an uncertain measure. Now we define an uncertain variable ξ as an identity function from the uncertainty space $(\Re, \mathcal{A}, \mathcal{M})$ to $\Re$. We may verify that ξ meets (1.38). The theorem is proved.

Theorem 1.15 *(Third Measure Inversion Theorem) Let ξ be an uncertain variable with third identification function (λ, ρ). Then for any Borel set B of*

real numbers, we have

$$\mathcal{M}\{\xi \in B\} = \begin{cases} \sup_{x \in B} \lambda(x) + \int_B \rho(x)\mathrm{d}x, \\ \qquad\qquad \textit{if } \sup_{x \in B} \lambda(x) + \int_B \rho(x)\mathrm{d}x < 0.5 \\ 1 - \sup_{x \in B^c} \lambda(x) - \int_{B^c} \rho(x)\mathrm{d}x, \\ \qquad\qquad \textit{if } \sup_{x \in B^c} \lambda(x) + \int_{B^c} \rho(x)\mathrm{d}x < 0.5 \\ 0.5, \qquad \textit{otherwise.} \end{cases}$$

Proof: It follows from the definition of third identification function immediately.

Example 1.15: Suppose ξ is a discrete uncertain variable taking values in $\{x_1, x_2, \cdots\}$ with *third identification coefficients* $\{(\lambda_1, \rho_1), (\lambda_2, \rho_2), \cdots\}$. Assume that

$$\sup_{x_i \in B} \lambda_i + \sum_{x_i \in B} \rho_i \geq 0.5 \quad \text{and/or} \quad \sup_{x_i \in B^c} \lambda_i + \sum_{x_i \in B^c} \rho_i \geq 0.5$$

for any set B of real numbers. Then we have

$$\mathcal{M}\{\xi \in B\} = \begin{cases} \sup_{x_i \in B} \lambda_i + \sum_{x_i \in B} \rho_i, & \text{if } \sup_{x_i \in B} \lambda_i + \sum_{x_i \in B} \rho_i < 0.5 \\ 1 - \sup_{x_i \in B^c} \lambda_i - \sum_{x_i \in B^c} \rho_i, & \text{if } \sup_{x_i \in B^c} \lambda_i + \sum_{x_i \in B^c} \rho_i < 0.5 \\ 0.5, & \text{otherwise.} \end{cases}$$

1.4   Uncertainty Distribution

This section introduces the concept of uncertainty distribution in order to describe uncertain variables. In many cases, it is sufficient to know the uncertainty distribution rather than the uncertain variable itself.

Definition 1.13 *(Liu [127]) The uncertainty distribution* $\Phi: \Re \to [0,1]$ *of an uncertain variable ξ is defined by*

$$\Phi(x) = \mathcal{M}\{\xi \leq x\}. \tag{1.39}$$

Theorem 1.16 *(Peng and Iwamura [186], Sufficient and Necessary Condition for Uncertainty Distribution) A function $\Phi : \Re \to [0,1]$ is an uncertainty distribution if and only if it is an increasing function except $\Phi(x) \equiv 0$ and $\Phi(x) \equiv 1$.*

Proof: It is obvious that an uncertainty distribution Φ is an increasing function. Conversely, suppose that Φ is an increasing function but $\Phi(x) \not\equiv 0$ and $\Phi(x) \not\equiv 1$. We will prove that there is an uncertain variable whose uncertainty distribution is just Φ. Let $\mathcal{C}$ be a collection of all intervals of the form $(-\infty, a]$, (b, ∞), $\emptyset$ and $\Re$. We define a set on $\Re$ as follows,

$$\mathcal{M}\{(-\infty, a]\} = \Phi(a),$$

$$\mathcal{M}\{(b, +\infty)\} = 1 - \Phi(b),$$

$$\mathcal{M}\{\emptyset\} = 0, \quad \mathcal{M}\{\Re\} = 1.$$

For an arbitrary Borel set B of real numbers, there exists a sequence $\{A_i\}$ in $\mathcal{C}$ such that

$$B \subset \bigcup_{i=1}^{\infty} A_i.$$

Note that such a sequence is not unique. Thus the set function $\mathcal{M}\{B\}$ is defined by

$$\mathcal{M}\{B\} = \begin{cases} \displaystyle\inf_{B \subset \cup A_i} \sum_{i=1}^{\infty} \mathcal{M}\{A_i\}, & \text{if } \displaystyle\inf_{B \subset \cup A_i} \sum_{i=1}^{\infty} \mathcal{M}\{A_i\} < 0.5 \\ 1 - \displaystyle\inf_{B^c \subset \cup A_i} \sum_{i=1}^{\infty} \mathcal{M}\{A_i\}, & \text{if } \displaystyle\inf_{B^c \subset \cup A_i} \sum_{i=1}^{\infty} \mathcal{M}\{A_i\} < 0.5 \\ 0.5, & \text{otherwise.} \end{cases}$$

We may prove that the set function $\mathcal{M}$ is indeed an uncertain measure on $\Re$, and the uncertain variable defined by the identity function $\xi(\gamma) = \gamma$ from the uncertainty space $(\Re, \mathcal{L}, \mathcal{M})$ to $\Re$ has the uncertainty distribution Φ.

Example 1.16: Let c be a number with $0 < c < 1$. Then $\Phi(x) \equiv c$ is an uncertainty distribution. When $c \leq 0.5$, we define a set function over $\Re$ as follows,

$$\mathcal{M}\{\Lambda\} = \begin{cases} 0, & \text{if } \Lambda = \emptyset \\ c, & \text{if } \Lambda \text{ is upper bounded} \\ 0.5, & \text{if both } \Lambda \text{ and } \Lambda^c \text{ are upper unbounded} \\ 1 - c, & \text{if } \Lambda^c \text{ is upper bounded} \\ 1, & \text{if } \Lambda = \Gamma. \end{cases}$$

Then $(\Re, \mathcal{L}, \mathcal{M})$ is an uncertainty space. It is easy to verify that the identity function $\xi(\gamma) = \gamma$ is an uncertain variable whose uncertainty distribution is

just $\Phi(x) \equiv c$. When $c > 0.5$, we define

$$\mathcal{M}\{\Lambda\} = \begin{cases} 0, & \text{if } \Lambda = \emptyset \\ 1 - c, & \text{if } \Lambda \text{ is upper bounded} \\ 0.5, & \text{if both } \Lambda \text{ and } \Lambda^c \text{ are upper unbounded} \\ c, & \text{if } \Lambda^c \text{ is upper bounded} \\ 1, & \text{if } \Lambda = \Gamma. \end{cases}$$

Then the function $\xi(\gamma) = -\gamma$ is an uncertain variable whose uncertainty distribution is just $\Phi(x) \equiv c$.

Example 1.17: Assume that two uncertain variables ξ and η have the same uncertainty distribution. One question is whether $\xi = \eta$ or not. Generally speaking, it is not true. Take $(\Gamma, \mathcal{L}, \mathcal{M})$ to be $\{\gamma_1, \gamma_2\}$ with

$$\mathcal{M}\{\gamma_1\} = \mathcal{M}\{\gamma_2\} = 0.5.$$

We now define two uncertain variables as follows,

$$\xi(\gamma) = \begin{cases} -1, & \text{if } \gamma = \gamma_1 \\ 1, & \text{if } \gamma = \gamma_2, \end{cases} \qquad \eta(\gamma) = \begin{cases} 1, & \text{if } \gamma = \gamma_1 \\ -1, & \text{if } \gamma = \gamma_2. \end{cases}$$

Then ξ and η have the same uncertainty distribution,

$$\Phi(x) = \begin{cases} 0, & \text{if } x < -1 \\ 0.5, & \text{if } -1 \le x < 1 \\ 1, & \text{if } x \ge 1. \end{cases}$$

However, it is clear that $\xi \ne \eta$ in the sense of Definition 1.6.

Definition 1.14 *An uncertain variable ξ is called linear if it has a linear uncertainty distribution*

$$\Phi(x) = \begin{cases} 0, & \text{if } x \le a \\ (x - a)/(b - a), & \text{if } a \le x \le b \\ 1, & \text{if } x \ge b \end{cases} \qquad (1.40)$$

denoted by $\mathcal{L}(a, b)$ where a and b are real numbers with $a < b$.

Definition 1.15 *An uncertain variable ξ is called zigzag if it has a zigzag uncertainty distribution*

$$\Phi(x) = \begin{cases} 0, & \text{if } x \le a \\ (x - a)/2(b - a), & \text{if } a \le x \le b \\ (x + c - 2b)/2(c - b), & \text{if } b \le x \le c \\ 1, & \text{if } x \ge c \end{cases} \qquad (1.41)$$

denoted by $\mathcal{Z}(a, b, c)$ where a, b, c are real numbers with $a < b < c$.
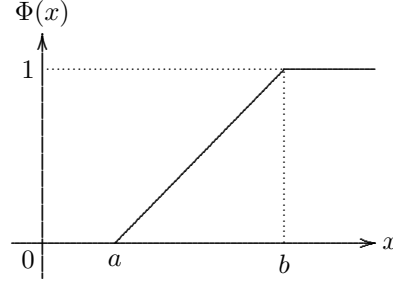
Figure 1.4: Linear Uncertainty Distribution
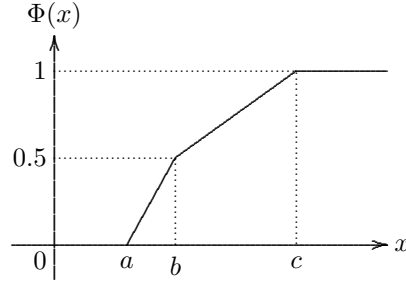
Figure 1.5: Zigzag Uncertainty Distribution

Definition 1.16 *An uncertain variable ξ is called normal if it has a normal uncertainty distribution*

$$\Phi(x) = \left(1 + \exp\left(\frac{\pi(e - x)}{\sqrt{3}\sigma}\right)\right)^{-1}, \quad x \in \Re \qquad (1.42)$$

denoted by $\mathcal{N}(e, \sigma)$ where e and σ are real numbers with $\sigma > 0$.

Definition 1.17 *An uncertain variable ξ is called lognormal if $\ln \xi$ is a normal uncertain variable $\mathcal{N}(e, \sigma)$. In other words, a lognormal uncertain variable has an uncertainty distribution*

$$\Phi(x) = \left(1 + \exp\left(\frac{\pi(e - \ln x)}{\sqrt{3}\sigma}\right)\right)^{-1}, \quad x \geq 0 \qquad (1.43)$$

denoted by $\mathcal{LOGN}(e, \sigma)$, where e and σ are real numbers with $\sigma > 0$.

Measure Inversion Theorem

Theorem 1.17 *(Measure Inversion Theorem) Let ξ be an uncertain variable with uncertainty distribution Φ. Then for any real number x, we have*

$$\mathcal{M}\{\xi \leq x\} = \Phi(x), \quad \mathcal{M}\{\xi > x\} = 1 - \Phi(x). \qquad (1.44)$$
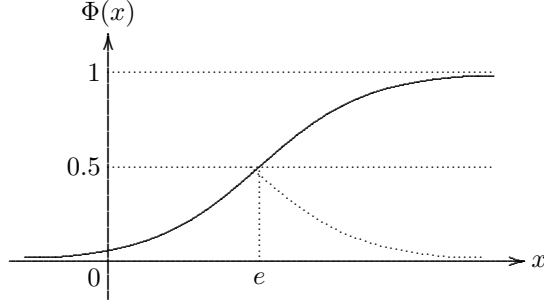
Figure 1.6: Normal Uncertainty Distribution
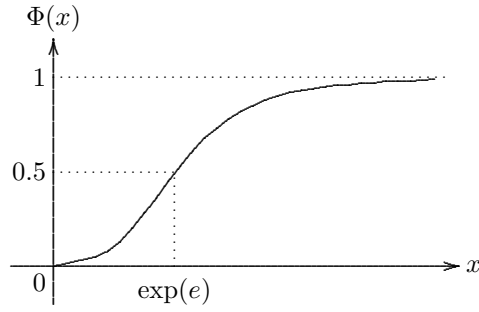
Figure 1.7: Lognormal Uncertainty Distribution

Proof: The equation $\mathcal{M}\{\xi \leq x\} = \Phi(x)$ follows from the definition of uncertainty distribution immediately. By using the self-duality of uncertain measure, we get $\mathcal{M}\{\xi > x\} = 1 - \mathcal{M}\{\xi \leq x\} = 1 - \Phi(x)$.

Remark 1.5: If the uncertainty distribution Φ is a continuous function, then we also have

$$\mathcal{M}\{\xi < x\} = \Phi(x), \quad \mathcal{M}\{\xi \geq x\} = 1 - \Phi(x). \tag{1.45}$$

Theorem 1.18 *Let ξ be an uncertain variable with continuous uncertainty distribution Φ. Then for any interval $[a, b]$, we have*

$$\Phi(b) - \Phi(a) \leq \mathcal{M}\{a \leq \xi \leq b\} \leq \Phi(b) \wedge (1 - \Phi(a)). \tag{1.46}$$

Proof: It follows from the subadditivity of uncertain measure and measure inversion theorem that

$$\mathcal{M}\{a \leq \xi \leq b\} + \mathcal{M}\{\xi \leq a\} \leq \mathcal{M}\{\xi \leq b\}.$$

That is,

$$\mathcal{M}\{a \leq \xi \leq b\} + \Phi(a) \leq \Phi(b).$$

Thus the inequality on the left hand side is verified. It follows from the monotonicity of uncertain measure and measure inversion theorem that

$$\mathcal{M}\{a \leq \xi \leq b\} \leq \mathcal{M}\{\xi \in (-\infty, b]\} = \Phi(b).$$

On the other hand,

$$\mathcal{M}\{a \leq x \leq b\} \leq \mathcal{M}\{\xi \in [a, +\infty)\} = 1 - \Phi(\inf a).$$

Hence the inequality on the right hand side is proved.

Remark 1.6: When a scalar is needed for $\mathcal{M}\{a \leq \xi \leq b\}$, it follows from the maximum uncertainty principle that

$$\mathcal{M}\{a \leq \xi \leq b\} = \begin{cases} \Phi(b) - \Phi(a), & \text{if } \Phi(b) - \Phi(a) > 0.5 \\ \Phi(b) \wedge (1 - \Phi(a)), & \text{if } \Phi(b) \wedge (1 - \Phi(a)) < 0.5 \\ 0.5, & \text{otherwise.} \end{cases} \quad (1.47)$$

Regular Uncertainty Distribution

Definition 1.18 *An uncertainty distribution Φ is said to be regular if its inverse function $\Phi^{-1}(\alpha)$ exists and is unique for each $\alpha \in (0,1)$.*

It is easy to verify that a regular uncertainty distribution Φ is a continuous function. In addition, Φ is strictly increasing at each point x with $0 < \Phi(x) < 1$. Furthermore,

$$\lim_{x \to -\infty} \Phi(x) = 0, \quad \lim_{x \to +\infty} \Phi(x) = 1. \quad (1.48)$$

For example, linear uncertainty distribution, zigzag uncertainty distribution, normal uncertainty distribution, and lognormal uncertainty distribution are all regular.

Joint Uncertainty Distribution

Definition 1.19 *Let $(\xi_1, \xi_2, \cdots, \xi_n)$ be an uncertain vector. Then the joint uncertainty distribution $\Phi : \Re^n \to [0,1]$ is defined by*

$$\Phi(x_1, x_2, \cdots, x_n) = \mathcal{M}\{\xi_1 \leq x_1, \xi_2 \leq x_2, \cdots, \xi_n \leq x_n\}. \quad (1.49)$$

1.5   Independence

Definition 1.20 *(Liu [130]) The uncertain variables $\xi_1, \xi_2, \cdots, \xi_m$ are said to be independent if*

$$\mathcal{M}\left\{\bigcap_{i=1}^{m}(\xi_i \in B_i)\right\} = \min_{1 \leq i \leq m} \mathcal{M}\{\xi_i \in B_i\} \quad (1.50)$$

for any Borel sets $B_1, B_2, \cdots, B_m$ of real numbers.

Example 1.18: Let ξ_1 be an uncertain variable and let ξ_2 be a constant c. For any Borel sets B_1 and B_2, if $c \in B_2$, then $\mathcal{M}\{\xi_2 \in B_2\} = 1$ and

$$\mathcal{M}\left\{(\xi_1 \in B_1) \cap (\xi_2 \in B_2)\right\} = \mathcal{M}\{\xi_1 \in B_1\} = \mathcal{M}\{\xi_1 \in B_1\} \wedge \mathcal{M}\{\xi_2 \in B_2\}.$$

If $c \notin B_2$, then $\mathcal{M}\{\xi_2 \in B_2\} = 0$ and

$$\mathcal{M}\left\{(\xi_1 \in B_1) \cap (\xi_2 \in B_2)\right\} = \mathcal{M}\{\emptyset\} = 0 = \mathcal{M}\{\xi_1 \in B_1\} \wedge \mathcal{M}\{\xi_2 \in B_2\}.$$

It follows from the definition of independence that an uncertain variable is always independent of a constant.

Theorem 1.19 *The uncertain variables $\xi_1, \xi_2, \cdots, \xi_m$ are independent if and only if*

$$\mathcal{M}\left\{\bigcup_{i=1}^{m}(\xi_i \in B_i)\right\} = \max_{1 \leq i \leq m} \mathcal{M}\left\{\xi_i \in B_i\right\} \tag{1.51}$$

for any Borel sets $B_1, B_2, \cdots, B_m$ of real numbers.

Proof: It follows from the self-duality of uncertain measure that $\xi_1, \xi_2, \cdots, \xi_m$ are independent if and only if

$$\mathcal{M}\left\{\bigcup_{i=1}^{m}(\xi_i \in B_i)\right\} = 1 - \mathcal{M}\left\{\bigcap_{i=1}^{m}(\xi_i \in B_i^c)\right\}$$
$$= 1 - \min_{1 \leq i \leq m} \mathcal{M}\{\xi_i \in B_i^c\} = \max_{1 \leq i \leq m} \mathcal{M}\left\{\xi_i \in B_i\right\}.$$

Thus the proof is complete.

Theorem 1.20 *Let Φ_i be uncertainty distributions of uncertain variables $\xi_i, \ i = 1, 2, \cdots, m$, respectively, and Φ the joint uncertainty distribution of uncertain vector $(\xi_1, \xi_2, \cdots, \xi_m)$. If $\xi_1, \xi_2, \cdots, \xi_m$ are independent, then we have*

$$\Phi(x_1, x_2, \cdots, x_m) = \min_{1 \leq i \leq m} \Phi_i(x_i) \tag{1.52}$$

for any real numbers $x_1, x_2, \cdots, x_m$.

Proof: Since $\xi_1, \xi_2, \cdots, \xi_m$ are independent uncertain variables, we have

$$\Phi(x_1, x_2, \cdots, x_m) = \mathcal{M}\left\{\bigcap_{i=1}^{m}(\xi_i \leq x_i)\right\} = \min_{1 \leq i \leq m} \mathcal{M}\{\xi_i \leq x_i\} = \min_{1 \leq i \leq m} \Phi_i(x_i)$$

for any real numbers $x_1, x_2, \cdots, x_m$. The theorem is proved.

Example 1.19: However, the equation (1.52) does not imply that the uncertain variables are independent. For example, let ξ be an uncertain variable

with uncertainty distribution Φ. Then the joint uncertainty distribution Ψ of uncertain vector (ξ, ξ) is

$$\Psi(x_1, x_2) = \mathcal{M}\{\xi \le x_1, \xi \le x_2\} = \mathcal{M}\{\xi \le x_1\} \wedge \mathcal{M}\{\xi \le x_2\} = \Phi(x_1) \wedge \Phi(x_2)$$

for any real numbers x_1 and x_2. But, generally speaking, an uncertain variable is not independent with itself.

Theorem 1.21 *Let $\xi_1, \xi_2, \cdots, \xi_m$ be independent uncertain variables, and $f_1, f_2, \cdots, f_n$ measurable functions. Then $f_1(\xi_1), f_2(\xi_2), \cdots, f_m(\xi_m)$ are independent uncertain variables.*

Proof: For any Borel sets $B_1, B_2, \cdots, B_m$ of $\Re$, it follows from the definition of independence that

$$\mathcal{M}\left\{\bigcap_{i=1}^{m}(f_i(\xi_i) \in B_i)\right\} = \mathcal{M}\left\{\bigcap_{i=1}^{m}(\xi_i \in f_i^{-1}(B_i))\right\}$$

$$= \min_{1 \le i \le m} \mathcal{M}\{\xi_i \in f_i^{-1}(B_i)\} = \min_{1 \le i \le m} \mathcal{M}\{f_i(\xi_i) \in B_i\}.$$

Thus $f_1(\xi_1), f_2(\xi_2), \cdots, f_m(\xi_m)$ are independent uncertain variables.

1.6   Operational Law

Theorem 1.22 *(Liu [130], Operational Law) Let $\xi_1, \xi_2, \cdots, \xi_n$ be independent uncertain variables, and $f : \Re^n \to \Re$ a measurable function. Then $\xi = f(\xi_1, \xi_2, \cdots, \xi_n)$ is an uncertain variable such that*

$$\mathcal{M}\{\xi \in B\} = \begin{cases} \sup_{f(B_1, B_2, \cdots, B_n) \subset B} \min_{1 \le k \le n} \mathcal{M}_k\{\xi_k \in B_k\}, \\ \qquad if \sup_{f(B_1, B_2, \cdots, B_n) \subset B} \min_{1 \le k \le n} \mathcal{M}_k\{\xi_k \in B_k\} > 0.5 \\ 1 - \sup_{f(B_1, B_2, \cdots, B_n) \subset B^c} \min_{1 \le k \le n} \mathcal{M}_k\{\xi_k \in B_k\}, \\ \qquad if \sup_{f(B_1, B_2, \cdots, B_n) \subset B^c} \min_{1 \le k \le n} \mathcal{M}_k\{\xi_k \in B_k\} > 0.5 \\ 0.5, \qquad otherwise \end{cases}$$

where $B, B_1, B_2, \cdots, B_n$ are Borel sets, and $f(B_1, B_2, \cdots, B_n) \subset B$ means $f(x_1, x_2, \cdots, x_n) \in B$ for any $x_1 \in B_1, x_2 \in B_2, \cdots, x_n \in B_n$.

Proof: Write $\Lambda = \{\xi \in B\}$ and $\Lambda_k = \{\xi_k \in B_k\}$ for $k = 1, 2, \cdots, n$. It is easy to verify that

$$\Lambda_1 \times \Lambda_2 \times \cdots \times \Lambda_n \subset \Lambda \text{ if and only if } f(B_1, B_2, \cdots, B_n) \subset B,$$

$$\Lambda_1 \times \Lambda_2 \times \cdots \times \Lambda_n \subset \Lambda^c \text{ if and only if } f(B_1, B_2, \cdots, B_n) \subset B^c.$$

Thus the operational law follows from the product measure axiom immediately.

Increasing Function of Single Uncertain Variable

Theorem 1.23 *Let ξ be an uncertain variable with regular uncertainty distribution Φ, and let f be a strictly increasing function. Then $f(\xi)$ has an uncertainty distribution*

$$\Psi(x) = \Phi(f^{-1}(x)), \quad x \in \Re \tag{1.53}$$

whose inverse function is

$$\Psi^{-1}(\alpha) = f(\Phi^{-1}(\alpha)), \quad 0 < \alpha < 1. \tag{1.54}$$

Proof: Since f is a strictly increasing function, we have, for each real number x,

$$f((-\infty, f^{-1}(x)]) = (-\infty, x].$$

It follows from the operational law that

$$\Psi(x) = \mathcal{M}\{f(\xi) \in (-\infty, x]\} = \mathcal{M}\left\{\xi \in (-\infty, f^{-1}(x)]\right\} = \Phi(f^{-1}(x)).$$

The first part is verified. Now let $\Psi(x) = \alpha$. Then

$$\Phi(f^{-1}(x)) = \alpha \quad \text{implies} \quad x = f(\Phi^{-1}(\alpha)).$$

Thus we have $\Psi^{-1}(\alpha) = x = f(\Phi^{-1}(\alpha))$. The theorem is proved.

99 Method 1.1 *It is suggested that an uncertain variable ξ with uncertainty distribution Φ is represented on a computer by*

$$\begin{array}{|c|c|c|c|c|}
\hline
0.01 & 0.02 & 0.03 & \cdots & 0.99 \\
\hline
x_1 & x_2 & x_3 & \cdots & x_{99} \\
\hline
\end{array} \tag{1.55}$$

where $0.01, 0.02, 0.03, \cdots, 0.99$ in the first row are the values of uncertainty distribution Φ, and $x_1, x_2, x_3, \cdots, x_{99}$ in the second row are the corresponding values of $\Phi^{-1}(0.01), \Phi^{-1}(0.02), \Phi^{-1}(0.03), \cdots, \Phi^{-1}(0.99)$. Essentially, it is a discrete representation of uncertainty distribution Φ. Then for any strictly increasing function $f(x)$, the uncertain variable $f(\xi)$ is

$$\begin{array}{|c|c|c|c|c|}
\hline
0.01 & 0.02 & 0.03 & \cdots & 0.99 \\
\hline
f(x_1) & f(x_2) & f(x_3) & \cdots & f(x_{99}) \\
\hline
\end{array} \tag{1.56}$$

The 99 method may be extended to the 999 method if a more precise result is needed.

Example 1.20: Let ξ be an uncertain variable with regular uncertainty distribution Φ. Then for any number $k > 0$, the uncertainty distribution of $k\xi$ is

$$\Psi(x) = \Phi\left(\frac{x}{k}\right), \quad x \in \Re \tag{1.57}$$

whose inverse function is

$$\Psi^{-1}(\alpha) = k\Phi^{-1}(\alpha), \quad 0 < \alpha < 1. \tag{1.58}$$

If ξ is represented by

0.01	0.02	0.03	$\cdots$	0.99
x_1	x_2	x_3	$\cdots$	x_{99}

$$\tag{1.59}$$

then the 99 method yields that $k\xi$ is

0.01	0.02	0.03	$\cdots$	0.99
kx_1	kx_2	kx_3	$\cdots$	kx_{99}

$$\tag{1.60}$$

Example 1.21: If ξ is an uncertain variables with regular uncertainty distribution Φ and k is a constant, then $\xi + k$ is an uncertain variable with uncertainty distribution

$$\Psi(x) = \Phi(x - k), \quad x \in \Re \tag{1.61}$$

whose inverse function is

$$\Psi^{-1}(\alpha) = \Phi^{-1}(\alpha) + k, \quad 0 < \alpha < 1. \tag{1.62}$$

If ξ is represented by

0.01	0.02	0.03	$\cdots$	0.99
x_1	x_2	x_3	$\cdots$	x_{99}

$$\tag{1.63}$$

then the 99 method yields that $\xi + k$ is

0.01	0.02	0.03	$\cdots$	0.99
$x_1 + k$	$x_2 + k$	$x_3 + k$	$\cdots$	$x_{99} + k$

$$\tag{1.64}$$

Example 1.22: Let ξ be a nonnegative uncertain variable with regular uncertainty distribution Φ. Since x^2 is a strictly increasing function on $[0, +\infty)$, the square ξ^2 is an uncertain variable with uncertainty distribution

$$\Psi(x) = \Phi(\sqrt{x}), \quad \forall x \geq 0 \tag{1.65}$$

whose inverse function is

$$\Psi^{-1}(\alpha) = \left(\Phi^{-1}(\alpha)\right)^2, \quad 0 < \alpha < 1. \tag{1.66}$$

If ξ is represented by

0.01	0.02	0.03	$\cdots$	0.99
x_1	x_2	x_3	$\cdots$	x_{99}

$$\tag{1.67}$$

then the 99 method yields that the uncertain variable ξ^2 is

$$
\begin{array}{|c|c|c|c|c|}
\hline
0.01 & 0.02 & 0.03 & \cdots & 0.99 \\
\hline
x_1^2 & x_2^2 & x_3^2 & \cdots & x_{99}^2 \\
\hline
\end{array}
\tag{1.68}
$$

Example 1.23: Let ξ be an uncertain variable with regular uncertainty distribution Φ. Since $\exp(x)$ is a strictly increasing function, $\exp(\xi)$ is an uncertain variable with uncertainty distribution

$$\Psi(x) = \Phi(\ln x), \quad \forall x > 0 \tag{1.69}$$

whose inverse function is

$$\Psi^{-1}(\alpha) = \exp\left(\Phi^{-1}(\alpha)\right), \quad 0 < \alpha < 1. \tag{1.70}$$

If ξ is represented by

$$
\begin{array}{|c|c|c|c|c|}
\hline
0.01 & 0.02 & 0.03 & \cdots & 0.99 \\
\hline
x_1 & x_2 & x_3 & \cdots & x_{99} \\
\hline
\end{array}
\tag{1.71}
$$

then the 99 method yields that the uncertain variable $\exp(\xi)$ is

$$
\begin{array}{|c|c|c|c|c|}
\hline
0.01 & 0.02 & 0.03 & \cdots & 0.99 \\
\hline
\exp(x_1) & \exp(x_2) & \exp(x_3) & \cdots & \exp(x_{99}) \\
\hline
\end{array}
\tag{1.72}
$$

Decreasing Function of Single Uncertain Variable

Theorem 1.24 *Let ξ be an uncertain variable with regular uncertainty distribution Φ, and let f be a strictly decreasing function. Then $f(\xi)$ has an uncertainty distribution*

$$\Psi(x) = 1 - \Phi(f^{-1}(x)), \quad x \in \Re \tag{1.73}$$

whose inverse function is

$$\Psi^{-1}(\alpha) = f(\Phi^{-1}(1-\alpha)), \quad 0 < \alpha < 1. \tag{1.74}$$

Proof: Since f is a strictly increasing function, we have, for each real number x,

$$f((-\infty, f^{-1}(x))) = (x, +\infty).$$

It follows from the operational law and continuity of Φ that

$$\Psi(x) = \mathcal{M}\{f(\xi) \in (-\infty, x]\} = 1 - \mathcal{M}\{f(\xi) \in (x, +\infty)\}$$

$$= 1 - \mathcal{M}\{\xi \in (-\infty, f^{-1}(x))\} = 1 - \Phi(f^{-1}(x)).$$

The first part is verified. Now let $\Psi(x) = \alpha$. Then

$$1 - \Phi(f^{-1}(x)) = \alpha \quad \text{implies} \quad x = f(\Phi^{-1}(1-\alpha)).$$

Thus we have $\Psi^{-1}(\alpha) = x = f(\Phi^{-1}(1-\alpha))$. The theorem is proved.

99 Method 1.2 *Let ξ be an uncertain variable represented on a computer by*

0.01	0.02	0.03	$\cdots$	0.99
x_1	x_2	x_3	$\cdots$	x_{99}

(1.75)

Then for any strictly decreasing function $f(x)$, the uncertain variable $f(\xi)$ is

0.01	0.02	0.03	$\cdots$	0.99
$f(x_{99})$	$f(x_{98})$	$f(x_{97})$	$\cdots$	$f(x_1)$

(1.76)

Example 1.24: Let ξ be an uncertain variable with regular uncertainty distribution Φ. Then $-\xi$ has an uncertainty distribution

$$\Psi(x) = 1 - \Phi(-x), \quad x \in \Re \tag{1.77}$$

whose inverse function is

$$\Psi^{-1}(\alpha) = -\Phi^{-1}(1 - \alpha), \quad 0 < \alpha < 1. \tag{1.78}$$

If ξ is represented by

0.01	0.02	0.03	$\cdots$	0.99
x_1	x_2	x_3	$\cdots$	x_{99}

(1.79)

then the 99 method yields that the uncertain variable $-\xi$ is

0.01	0.02	0.03	$\cdots$	0.99
$-x_{99}$	$-x_{98}$	$-x_{97}$	$\cdots$	$-x_1$

(1.80)

Example 1.25: Let ξ be a positive uncertain variable with regular uncertainty distribution Φ. Since $1/x$ is a strictly decreasing function on $(0, +\infty)$, the reciprocal $1/\xi$ is an uncertain variable with uncertainty distribution

$$\Psi(x) = 1 - \Phi\left(\frac{1}{x}\right), \quad \forall x > 0 \tag{1.81}$$

whose inverse function is

$$\Psi^{-1}(\alpha) = \frac{1}{\Phi^{-1}(1 - \alpha)}, \quad 0 < \alpha < 1. \tag{1.82}$$

If ξ is represented by

0.01	0.02	0.03	$\cdots$	0.99
x_1	x_2	x_3	$\cdots$	x_{99}

(1.83)

then the 99 method yields that the uncertain variable $1/\xi$ is

0.01	0.02	0.03	$\cdots$	0.99
$1/x_{99}$	$1/x_{98}$	$1/x_{97}$	$\cdots$	$1/x_1$

(1.84)

Example 1.26: Let ξ be an uncertain variable with regular uncertainty distribution Φ. Since $\exp(-x)$ is a strictly decreasing function, $\exp(-\xi)$ is an uncertain variable with uncertainty distribution

$$\Psi(x) = \Phi(-\ln x), \quad \forall x > 0 \tag{1.85}$$

whose inverse function is

$$\Psi^{-1}(\alpha) = \exp\left(-\Phi^{-1}(1-\alpha)\right), \quad 0 < \alpha < 1. \tag{1.86}$$

If ξ is represented by

0.01	0.02	0.03	$\cdots$	0.99
x_1	x_2	x_3	$\cdots$	x_{99}

$$\tag{1.87}$$

then the 99 method yields that the uncertain variable $\exp(-\xi)$ is

0.01	0.02	0.03	$\cdots$	0.99
$\exp(-x_{99})$	$\exp(-x_{98})$	$\exp(-x_{97})$	$\cdots$	$\exp(-x_1)$

$$\tag{1.88}$$

Increasing Function of Multiple Uncertain Variables

Theorem 1.25 *Let $\xi_1, \xi_2, \cdots, \xi_n$ be independent uncertain variables with regular uncertainty distributions $\Phi_1, \Phi_2, \cdots, \Phi_n$, respectively. If $f : \Re^n \to \Re$ is a strictly increasing function, then*

$$\xi = f(\xi_1, \xi_2, \cdots, \xi_n) \tag{1.89}$$

is an uncertain variable with uncertainty distribution

$$\Psi(x) = \sup_{f(x_1, x_2, \cdots, x_n) = x} \min_{1 \leq i \leq n} \Phi_i(x_i), \quad x \in \Re \tag{1.90}$$

whose inverse function is

$$\Psi^{-1}(\alpha) = f(\Phi_1^{-1}(\alpha), \Phi_2^{-1}(\alpha), \cdots, \Phi_n^{-1}(\alpha)), \quad 0 < \alpha < 1. \tag{1.91}$$

Proof: Since $\xi_1, \xi_2, \cdots, \xi_n$ are independent uncertain variables and f is a strictly increasing function, we have

$$\mathcal{M}\{\xi \leq \Psi^{-1}(\alpha)\}$$
$$= \mathcal{M}\{f(\xi_1, \xi_2, \cdots, \xi_n) \leq f(\Phi_1^{-1}(\alpha), \Phi_2^{-1}(\alpha), \cdots, \Phi_n^{-1}(\alpha))\}$$
$$\geq \mathcal{M}\{(\xi_1 \leq \Phi_1^{-1}(\alpha)) \cap (\xi_2 \leq \Phi_2^{-1}(\alpha)) \cap \cdots \cap (\xi_n \leq \Phi_n^{-1}(\alpha))\}$$
$$= \mathcal{M}\{\xi_1 \leq \Phi_1^{-1}(\alpha)\} \wedge \mathcal{M}\{\xi_2 \leq \Phi_2^{-1}(\alpha)\} \wedge \cdots \wedge \mathcal{M}\{\xi_n \leq \Phi_n^{-1}(\alpha)\}$$
$$= \alpha \wedge \alpha \wedge \cdots \wedge \alpha = \alpha.$$

On the other hand, there exists some index i such that

$$\{f(\xi_1, \xi_2, \cdots, \xi_n) \leq f(\Phi_1^{-1}(\alpha), \Phi_2^{-1}(\alpha), \cdots, \Phi_n^{-1}(\alpha))\} \subset \{\xi_i \leq \Phi_i^{-1}(\alpha)\}.$$

Thus

$$\mathcal{M}\{\xi \leq \Psi^{-1}(\alpha)\} \leq \mathcal{M}\{\xi_i \leq \Phi_i^{-1}(\alpha)\} = \alpha.$$

It follows that $\mathcal{M}\{\xi \leq \Psi^{-1}(\alpha)\} = \alpha$. In other words, Ψ is just the uncertainty distribution of ξ. The theorem is proved.

99 Method 1.3 *Assume $\xi_1, \xi_2, \cdots, \xi_n$ are uncertain variables, and each ξ_i is represented on a computer by*

$$
\begin{array}{|c|c|c|c|c|}
\hline
0.01 & 0.02 & 0.03 & \cdots & 0.99 \\
\hline
x_1^i & x_2^i & x_3^i & \cdots & x_{99}^i \\
\hline
\end{array}
\tag{1.92}
$$

Then for any strictly increasing function $f(x_1, x_2, \cdots, x_n)$, the uncertain variable $f(\xi_1, \xi_2, \cdots, \xi_n)$ is

$$
\begin{array}{|c|c|c|c|}
\hline
0.01 & 0.02 & \cdots & 0.99 \\
\hline
f(x_1^1, x_1^2, \cdots, x_1^n) & f(x_2^1, x_2^2, \cdots, x_2^n) & \cdots & f(x_{99}^1, x_{99}^2, \cdots, x_{99}^n) \\
\hline
\end{array}
\tag{1.93}
$$

Example 1.27: Let $\xi_1, \xi_2, \cdots, \xi_n$ be independent uncertain variables with regular uncertainty distributions $\Phi_1, \Phi_2, \cdots, \Phi_n$, respectively. Then the sum

$$\xi = \xi_1 + \xi_2 + \cdots + \xi_n \tag{1.94}$$

is an uncertain variable with uncertainty distribution

$$\Psi(x) = \sup_{x_1 + x_2 + \cdots + x_n = x} \min_{1 \leq i \leq n} \Phi_i(x_i), \quad x \in \Re \tag{1.95}$$

whose inverse function is

$$\Psi^{-1}(\alpha) = \Phi_1^{-1}(\alpha) + \Phi_2^{-1}(\alpha) + \cdots + \Phi_n^{-1}(\alpha), \quad 0 < \alpha < 1. \tag{1.96}$$

If each ξ_i $(1 \leq i \leq n)$ is represented by

$$
\begin{array}{|c|c|c|c|c|}
\hline
0.01 & 0.02 & 0.03 & \cdots & 0.99 \\
\hline
x_1^i & x_2^i & x_3^i & \cdots & x_{99}^i \\
\hline
\end{array}
\tag{1.97}
$$

then the 99 method yields that the sum $\xi_1 + \xi_2 + \cdots + \xi_n$ is

$$
\begin{array}{|c|c|c|c|c|}
\hline
0.01 & 0.02 & 0.03 & \cdots & 0.99 \\
\hline
\displaystyle\sum_{i=1}^{n} x_1^i & \displaystyle\sum_{i=1}^{n} x_2^i & \displaystyle\sum_{i=1}^{n} x_3^i & \cdots & \displaystyle\sum_{i=1}^{n} x_{99}^i \\
\hline
\end{array}
\tag{1.98}
$$

Example 1.28: Let $\xi_1, \xi_2, \cdots, \xi_n$ be independent and positive uncertain variables with regular uncertainty distributions $\Phi_1, \Phi_2, \cdots, \Phi_n$, respectively. Then the product

$$\xi = \xi_1 \times \xi_2 \times \cdots \times \xi_n \tag{1.99}$$

is an uncertain variable with uncertainty distribution

$$\Psi(x) = \sup_{x_1 \times x_2 \times \cdots \times x_n = x} \min_{1 \leq i \leq n} \Phi_i(x_i), \quad x \in \Re \tag{1.100}$$

whose inverse function is

$$\Psi^{-1}(\alpha) = \Phi_1^{-1}(\alpha) \times \Phi_2^{-1}(\alpha) \times \cdots \times \Phi_n^{-1}(\alpha), \quad 0 < \alpha < 1. \tag{1.101}$$

If each ξ_i $(1 \leq i \leq n)$ is represented by

$$\begin{array}{|c|c|c|c|c|} \hline 0.01 & 0.02 & 0.03 & \cdots & 0.99 \\ \hline x_1^i & x_2^i & x_3^i & \cdots & x_{99}^i \\ \hline \end{array} \tag{1.102}$$

then the 99 method yileds that the product $\xi_1 \times \xi_2 \times \cdots \times \xi_n$ is

$$\begin{array}{|c|c|c|c|c|} \hline 0.01 & 0.02 & 0.03 & \cdots & 0.99 \\ \hline \prod_{i=1}^{n} x_1^i & \prod_{i=1}^{n} x_2^i & \prod_{i=1}^{n} x_3^i & \cdots & \prod_{i=1}^{n} x_{99}^i \\ \hline \end{array} \tag{1.103}$$

Example 1.29: Assume ξ_1, ξ_2, ξ_3 are independent and nonnegative uncertain variables with regular uncertainty distributions Φ_1, Φ_2, Φ_3, respectively. Then the uncertainty distribution of $(\xi_1 + \xi_2)\xi_3$ is

$$\Psi(x) = \sup_{(x_1+x_2)x_3=x} \Phi_1(x_1) \wedge \Phi_2(x_2) \wedge \Phi_3(x_3), \quad x \geq 0 \tag{1.104}$$

whose inverse function is

$$\Psi^{-1}(\alpha) = \left(\Phi_1^{-1}(\alpha) + \Phi_2^{-1}(\alpha) \right) \Phi_3^{-1}(\alpha), \quad 0 < \alpha < 1. \tag{1.105}$$

If ξ_1, ξ_2, ξ_3 are respectively represented by

$$\begin{array}{|c|c|c|c|c|} \hline 0.01 & 0.02 & 0.03 & \cdots & 0.99 \\ \hline x_1^1 & x_2^1 & x_3^1 & \cdots & x_{99}^1 \\ \hline \end{array}$$
$$\begin{array}{|c|c|c|c|c|} \hline 0.01 & 0.02 & 0.03 & \cdots & 0.99 \\ \hline x_1^2 & x_2^2 & x_3^2 & \cdots & x_{99}^2 \\ \hline \end{array} \tag{1.106}$$
$$\begin{array}{|c|c|c|c|c|} \hline 0.01 & 0.02 & 0.03 & \cdots & 0.99 \\ \hline x_1^3 & x_2^3 & x_3^3 & \cdots & x_{99}^3 \\ \hline \end{array}$$

then the 99 method yields that the uncertain variable $(\xi_1 + \xi_2)\xi_3$ is

$$\begin{array}{|c|c|c|c|} \hline 0.01 & 0.02 & \cdots & 0.99 \\ \hline (x_1^1 + x_1^2)x_1^3 & (x_2^1 + x_2^2)x_2^3 & \cdots & (x_{99}^1 + x_{99}^2)x_{99}^3 \\ \hline \end{array} \tag{1.107}$$

Theorem 1.26 *Assume that ξ_1 and ξ_2 are independent linear uncertain variables $\mathcal{L}(a_1, b_1)$ and $\mathcal{L}(a_2, b_2)$, respectively. Then the sum $\xi_1 + \xi_2$ is also a linear uncertain variable $\mathcal{L}(a_1 + a_2, b_1 + b_2)$, i.e.,*

$$\mathcal{L}(a_1, b_1) + \mathcal{L}(a_2, b_2) = \mathcal{L}(a_1 + a_2, b_1 + b_2). \tag{1.108}$$

The product of a linear uncertain variable $\mathcal{L}(a, b)$ and a scalar number $k > 0$ is also a linear uncertain variable $\mathcal{L}(ka, kb)$, i.e.,

$$k \cdot \mathcal{L}(a, b) = \mathcal{L}(ka, kb). \tag{1.109}$$

Proof: Assume that the uncertain variables ξ_1 and ξ_2 have uncertainty distributions Φ_1 and Φ_2, respectively. Then

$$\Phi_1^{-1}(\alpha) = (1 - \alpha)a_1 + \alpha b_1,$$

$$\Phi_2^{-1}(\alpha) = (1 - \alpha)a_2 + \alpha b_2.$$

It follows from the operational law that the uncertainty distribution of $\xi_1 + \xi_2$ is described by

$$\Psi^{-1}(\alpha) = \Phi_1^{-1}(\alpha) + \Phi_2^{-1}(\alpha) = (1 - \alpha)(a_1 + a_2) + \alpha(b_1 + b_2).$$

Hence the sum is also a linear uncertain variable $\mathcal{L}(a_1 + a_2, b_1 + b_2)$. The first part is verified. Next, suppose that the uncertainty distribution of the uncertain variable $\xi \sim \mathcal{L}(a, b)$ is Φ. It follows from the operational law that when $k > 0$, the uncertainty distribution of $k\xi$ is described by

$$\Psi^{-1}(\alpha) = k\Phi^{-1}(\alpha) = (1 - \alpha)(ka) + \alpha(kb).$$

Hence $k\xi$ is just a linear uncertain variable $\mathcal{L}(ka, kb)$.

Theorem 1.27 *Assume that ξ_1 and ξ_2 are independent zigzag uncertain variables $\mathcal{Z}(a_1, b_1, c_1)$ and $\mathcal{Z}(a_2, b_2, c_3)$, respectively. Then the sum $\xi_1 + \xi_2$ is also a zigzag uncertain variable $\mathcal{Z}(a_1 + a_2, b_1 + b_2, c_1 + c_2)$, i.e.,*

$$\mathcal{Z}(a_1, b_1, c_1) + \mathcal{Z}(a_2, b_2, c_2) = \mathcal{Z}(a_1 + a_2, b_1 + b_2, c_1 + c_2). \tag{1.110}$$

The product of a zigzag uncertain variable $\mathcal{Z}(a, b, c)$ and a scalar number $k > 0$ is also a zigzag uncertain variable $\mathcal{Z}(ka, kb, kc)$, i.e.,

$$k \cdot \mathcal{Z}(a, b, c) = \mathcal{Z}(ka, kb, kc). \tag{1.111}$$

Proof: Assume that the uncertain variables ξ_1 and ξ_2 have uncertainty distributions Φ_1 and Φ_2, respectively. Then

$$\Phi_1^{-1}(\alpha) = \begin{cases} (1 - 2\alpha)a_1 + 2\alpha b_1, & \text{if } \alpha < 0.5 \\ (2 - 2\alpha)b_1 + (2\alpha - 1)c_1, & \text{if } \alpha \geq 0.5, \end{cases}$$

$$\Phi_2^{-1}(\alpha) = \begin{cases} (1 - 2\alpha)a_2 + 2\alpha b_2, & \text{if } \alpha < 0.5 \\ (2 - 2\alpha)b_2 + (2\alpha - 1)c_2, & \text{if } \alpha \geq 0.5. \end{cases}$$

It follows from the operational law that the uncertainty distribution of $\xi_1 + \xi_2$ is described by

$$\Psi^{-1}(\alpha) = \begin{cases} (1 - 2\alpha)(a_1 + a_2) + 2\alpha(b_1 + b_2), & \text{if } \alpha < 0.5 \\ (2 - 2\alpha)(b_1 + b_2) + (2\alpha - 1)(c_1 + c_2), & \text{if } \alpha \geq 0.5. \end{cases}$$

Hence the sum is also a zigzag uncertain variable $\mathcal{Z}(a_1 + a_2, b_1 + b_2, c_1 + c_2)$. The first part is verified. Next, suppose that the uncertainty distribution of the uncertain variable $\xi \sim \mathcal{Z}(a, b, c)$ is Φ. It follows from the operational law that when $k > 0$, the uncertainty distribution of $k\xi$ is described by

$$\Psi^{-1}(\alpha) = k\Phi^{-1}(\alpha) = \begin{cases} (1 - 2\alpha)(ka) + 2\alpha(kb), & \text{if } \alpha < 0.5 \\ (2 - 2\alpha)(kb) + (2\alpha - 1)(kc), & \text{if } \alpha \geq 0.5. \end{cases}$$

Hence $k\xi$ is just a zigzag uncertain variable $\mathcal{Z}(ka, kb, kc)$.

Theorem 1.28 *Let ξ_1 and ξ_2 be independent normal uncertain variables $\mathcal{N}(e_1, \sigma_1)$ and $\mathcal{N}(e_2, \sigma_2)$, respectively. Then the sum $\xi_1 + \xi_2$ is also a normal uncertain variable $\mathcal{N}(e_1 + e_2, \sigma_1 + \sigma_2)$, i.e.,*

$$\mathcal{N}(e_1, \sigma_1) + \mathcal{N}(e_2, \sigma_2) = \mathcal{N}(e_1 + e_2, \sigma_1 + \sigma_2). \tag{1.112}$$

The product of a normal uncertain variable $\mathcal{N}(e, \sigma)$ and a scalar number $k > 0$ is also a normal uncertain variable $\mathcal{N}(ke, k\sigma)$, i.e.,

$$k \cdot \mathcal{N}(e, \sigma) = \mathcal{N}(ke, k\sigma). \tag{1.113}$$

Proof: Assume that the uncertain variables ξ_1 and ξ_2 have uncertainty distributions Φ_1 and Φ_2, respectively. Then

$$\Phi_1^{-1}(\alpha) = e_1 + \frac{\sigma_1\sqrt{3}}{\pi} \ln \frac{\alpha}{1 - \alpha},$$

$$\Phi_2^{-1}(\alpha) = e_2 + \frac{\sigma_2\sqrt{3}}{\pi} \ln \frac{\alpha}{1 - \alpha}.$$

It follows from the operational law that the uncertainty distribution of $\xi_1 + \xi_2$ is described by

$$\Psi^{-1}(\alpha) = \Phi_1^{-1}(\alpha) + \Phi_2^{-1}(\alpha) = (e_1 + e_2) + \frac{(\sigma_1 + \sigma_2)\sqrt{3}}{\pi} \ln \frac{\alpha}{1 - \alpha}.$$

Hence the sum is also a normal uncertain variable $\mathcal{N}(e_1 + e_2, \sigma_1 + \sigma_2)$. The first part is verified. Next, suppose that the uncertainty distribution of the

uncertain variable $\xi \sim \mathcal{N}(e, \sigma)$ is Φ. It follows from the operational law that, when $k > 0$, the uncertainty distribution of $k\xi$ is described by

$$\Psi^{-1}(\alpha) = k\Phi^{-1}(\alpha) = (ke) + \frac{(k\sigma)\sqrt{3}}{\pi} \ln \frac{\alpha}{1-\alpha}.$$

Hence $k\xi$ is just a normal uncertain variable $\mathcal{N}(ke, k\sigma)$.

Theorem 1.29 *Assume that $\xi_1, \xi_2, \cdots, \xi_n$ are independent lognormal uncertain variables $\mathcal{LOGN}(e_1, \sigma_1), \mathcal{LOGN}(e_2, \sigma_2), \cdots, \mathcal{LOGN}(e_n, \sigma_n)$, respectively. Then the sum $\xi_1 + \xi_2 + \cdots + \xi_n$ is an uncertain variable (not lognormal) whose uncertainty distribution Ψ is described by*

$$\Psi^{-1}(\alpha) = \sum_{i=1}^{n} \exp(e_i) \left(\frac{\alpha}{1-\alpha} \right)^{\sqrt{3}\sigma_i/\pi}, \quad 0 < \alpha < 1. \tag{1.114}$$

The product of a lognormal uncertain variable $\mathcal{LOGN}(e, \sigma)$ and a scalar number $k > 0$ is also a lognormal uncertain variable $\mathcal{LOGN}(e + \ln k, \sigma)$, i.e.,

$$k \cdot \mathcal{LOGN}(e, \sigma) = \mathcal{LOGN}(e + \ln k, \sigma). \tag{1.115}$$

Proof: Assume that the uncertain variables $\xi_1, \xi_2, \cdots, \xi_n$ have uncertainty distributions $\Phi_1, \Phi_2, \cdots, \Phi_n$, respectively. Then

$$\Phi_i^{-1}(\alpha) = \exp(e_i) \left(\frac{\alpha}{1-\alpha} \right)^{\sqrt{3}\sigma_i/\pi}$$

for $i = 1, 2, \cdots, n$. It follows from the operational law that the uncertainty distribution of $\xi_1 + \xi_2 + \cdots + \xi_n$ is described by (1.114). The first part is verified. Next, suppose that the uncertainty distribution of the uncertain variable $\xi \sim \mathcal{LOGN}(e, \sigma)$ is Φ. It follows from the operational law that, when $k > 0$, the uncertainty distribution of $k\xi$ is described by

$$\Psi^{-1}(\alpha) = k\Phi^{-1}(\alpha) = \exp(e + \ln k) \left(\frac{\alpha}{1-\alpha} \right)^{\sqrt{3}\sigma/\pi}.$$

Hence $k\xi$ is just a lognormal uncertain variable $\mathcal{LOGN}(e + \ln k, \sigma)$.

Example 1.30: Let $\xi_1, \xi_2, \cdots, \xi_n$ be independent uncertain variables with regular uncertainty distributions $\Phi_1, \Phi_2, \cdots, \Phi_n$, respectively. Then the maximum

$$\xi = \xi_1 \vee \xi_2 \vee \cdots \vee \xi_n \tag{1.116}$$

is an uncertain variable with uncertainty distribution

$$\Psi(x) = \Phi_1(x) \wedge \Phi_2(x) \wedge \cdots \wedge \Phi_n(x), \quad x \in \Re \tag{1.117}$$

whose inverse function is

$$\Psi^{-1}(\alpha) = \Phi_1^{-1}(\alpha) \vee \Phi_2^{-1}(\alpha) \vee \cdots \vee \Phi_n^{-1}(\alpha), \quad 0 < \alpha < 1. \qquad (1.118)$$

If each ξ_i $(1 \leq i \leq n)$ is represented by

0.01	0.02	0.03	$\cdots$	0.99
x_1^i	x_2^i	x_3^i	$\cdots$	x_{99}^i

$$(1.119)$$

then the 99 method yields that the maximum $\xi_1 \vee \xi_2 \vee \cdots \vee \xi_n$ is

0.01	0.02	0.03	$\cdots$	0.99
$\displaystyle\bigvee_{i=1}^{n} x_1^i$	$\displaystyle\bigvee_{i=1}^{n} x_2^i$	$\displaystyle\bigvee_{i=1}^{n} x_3^i$	$\cdots$	$\displaystyle\bigvee_{i=1}^{n} x_{99}^i$

$$(1.120)$$

Example 1.31: Let $\xi_1, \xi_2, \cdots, \xi_n$ be independent uncertain variables with regular uncertainty distributions $\Phi_1, \Phi_2, \cdots, \Phi_n$, respectively. Then the minimum

$$\xi = \xi_1 \wedge \xi_2 \wedge \cdots \wedge \xi_n \qquad (1.121)$$

is an uncertain variable with uncertainty distribution

$$\Psi(x) = \Phi_1(x) \vee \Phi_2(x) \vee \cdots \vee \Phi_n(x), \quad x \in \Re \qquad (1.122)$$

whose inverse function is

$$\Psi^{-1}(\alpha) = \Phi_1^{-1}(\alpha) \wedge \Phi_2^{-1}(\alpha) \wedge \cdots \wedge \Phi_n^{-1}(\alpha), \quad 0 < \alpha < 1. \qquad (1.123)$$

If each ξ_i $(1 \leq i \leq n)$ is represented by

0.01	0.02	0.03	$\cdots$	0.99
x_1^i	x_2^i	x_3^i	$\cdots$	x_{99}^i

$$(1.124)$$

then the 99 method yields that the minimum $\xi_1 \wedge \xi_2 \wedge \cdots \wedge \xi_n$ is

0.01	0.02	0.03	$\cdots$	0.99
$\displaystyle\bigwedge_{i=1}^{n} x_1^i$	$\displaystyle\bigwedge_{i=1}^{n} x_2^i$	$\displaystyle\bigwedge_{i=1}^{n} x_3^i$	$\cdots$	$\displaystyle\bigwedge_{i=1}^{n} x_{99}^i$

$$(1.125)$$

Example 1.32: If ξ is an uncertain variable with regular uncertainty distribution Φ and k is a constant, then $\xi \vee k$ is an uncertain variable with uncertainty distribution

$$\Psi(x) = \begin{cases} 0, & \text{if } x < k \\ \Phi(x), & \text{if } x \geq k \end{cases} \qquad (1.126)$$

and is produced by the 99 method as follows,

0.01	0.02	0.03	$\cdots$	0.99
$x_1 \vee k$	$x_2 \vee k$	$x_3 \vee k$	$\cdots$	$x_{99} \vee k$

$$(1.127)$$

In addition, $\xi \wedge k$ is an uncertain variable with uncertainty distribution

$$\Psi(x) = \begin{cases} \Phi(x), & \text{if } x < k \\ 1, & \text{if } x \geq k \end{cases} \qquad (1.128)$$

and is produced by the 99 method as follows,

0.01	0.02	0.03	$\cdots$	0.99
$x_1 \wedge k$	$x_2 \wedge k$	$x_3 \wedge k$	$\cdots$	$x_{99} \wedge k$

$$(1.129)$$

Decreasing Function of Multiple Uncertain Variables

Theorem 1.30 *Let $\xi_1, \xi_2, \cdots, \xi_n$ be independent uncertain variables with regular uncertainty distributions $\Phi_1, \Phi_2, \cdots, \Phi_n$, respectively. If $f : \Re^n \to \Re$ is a strictly decreasing function, then*

$$\xi = f(\xi_1, \xi_2, \cdots, \xi_n) \qquad (1.130)$$

is an uncertain variable with uncertainty distribution

$$\Psi(x) = \sup_{f(x_1, x_2, \cdots, x_n) = x} \min_{1 \leq i \leq n} (1 - \Phi_i(x_i)), \quad x \in \Re \qquad (1.131)$$

whose inverse function is

$$\Psi^{-1}(\alpha) = f(\Phi_1^{-1}(1-\alpha), \Phi_2^{-1}(1-\alpha), \cdots, \Phi_n^{-1}(1-\alpha)), \quad 0 < \alpha < 1. \quad (1.132)$$

Proof: Since $\xi_1, \xi_2, \cdots, \xi_n$ are independent uncertain variables and f is a strictly decreasing function, we have

$$\mathcal{M}\{\xi \leq \Psi^{-1}(\alpha)\}$$
$$= \mathcal{M}\{f(\xi_1, \xi_2, \cdots, \xi_n) \leq f(\Phi_1^{-1}(1-\alpha), \Phi_2^{-1}(1-\alpha), \cdots, \Phi_n^{-1}(1-\alpha))\}$$
$$\geq \mathcal{M}\{(\xi_1 \geq \Phi_1^{-1}(1-\alpha)) \cap (\xi_2 \geq \Phi_2^{-1}(1-\alpha)) \cap \cdots \cap (\xi_n \geq \Phi_n^{-1}(1-\alpha))\}$$
$$= \mathcal{M}\{\xi_1 \geq \Phi_1^{-1}(1-\alpha)\} \wedge \mathcal{M}\{\xi_2 \geq \Phi_2^{-1}(1-\alpha)\} \wedge \cdots \wedge \mathcal{M}\{\xi_n \geq \Phi_n^{-1}(1-\alpha)\}$$
$$= \alpha \wedge \alpha \wedge \cdots \wedge \alpha = \alpha. \quad \text{(By the continuity of } \Phi_i\text{'s)}$$

On the other hand, there exists some index i such that

$$\{f(\xi_1, \cdots, \xi_n) \leq f(\Phi_1^{-1}(1-\alpha), \cdots, \Phi_n^{-1}(1-\alpha))\} \subset \{\xi_i \geq \Phi_i^{-1}(1-\alpha)\}.$$

Thus

$$\mathcal{M}\{\xi \leq \Psi^{-1}(\alpha)\} \leq \mathcal{M}\{\xi_i \geq \Phi_i^{-1}(1-\alpha)\} = \alpha.$$

It follows that $\mathcal{M}\{\xi \leq \Psi^{-1}(\alpha)\} = \alpha$. In other words, Ψ is just the uncertainty distribution of ξ. The theorem is proved.

99 Method 1.4 *Assume $\xi_1, \xi_2, \cdots, \xi_n$ are uncertain variables, and each ξ_i is represented on a computer by*

0.01	0.02	0.03	$\cdots$	0.99
x_1^i	x_2^i	x_3^i	$\cdots$	x_{99}^i

$$(1.133)$$

Then for any strictly decreasing function $f(x_1, x_2, \cdots, x_n)$, the uncertain variable $f(\xi_1, \xi_2, \cdots, \xi_n)$ is

0.01	0.02	$\cdots$	0.99
$f(x_{99}^1, x_{99}^2, \cdots, x_{99}^n)$	$f(x_{98}^1, x_{98}^2, \cdots, x_{98}^n)$	$\cdots$	$f(x_1^1, x_1^2, \cdots, x_1^n)$

Alternating Monotone Function of Multiple Uncertain Variables

Theorem 1.31 *Let $\xi_1, \xi_2, \cdots, \xi_n$ be independent uncertain variables with regular uncertainty distributions $\Phi_1, \Phi_2, \cdots, \Phi_n$, respectively. If the function $f(x_1, x_2, \cdots, x_n)$ is strictly increasing with respect to $x_1, x_2, \cdots, x_m$ and strictly decreasing with respect to $x_{m+1}, x_{m+2}, \cdots, x_n$, then*

$$\xi = f(\xi_1, \xi_2, \cdots, \xi_n)$$

is an uncertain variable with uncertainty distribution

$$\Psi(x) = \sup_{f(x_1, x_2, \cdots, x_n)=x} \left(\min_{1 \leq i \leq m} \Phi_i(x_i) \wedge \min_{m+1 \leq i \leq n} (1 - \Phi_i(x_i)) \right), \quad x \in \Re$$

whose inverse function is

$$\Psi^{-1}(\alpha) = f(\Phi_1^{-1}(\alpha), \cdots, \Phi_m^{-1}(\alpha), \Phi_{m+1}^{-1}(1-\alpha), \cdots, \Phi_n^{-1}(1-\alpha)), \quad 0 < \alpha < 1.$$

Proof: We only prove the case of $m = 1$ and $n = 2$. Since ξ_1 and ξ_2 are independent uncertain variables and the function $f(x_1, x_2)$ is strictly increasing with respect to x_1 and strictly decreasing with x_2, we have

$$\begin{aligned}
\mathcal{M}\{\xi \leq \Psi^{-1}(\alpha)\} &= \mathcal{M}\{f(\xi_1, \xi_2) \leq f(\Phi_1^{-1}(\alpha), \Phi_2^{-1}(1-\alpha))\} \\
&\geq \mathcal{M}\{(\xi_1 \leq \Phi_1^{-1}(\alpha)) \cap (\xi_2 \geq \Phi_2^{-1}(1-\alpha))\} \\
&= \mathcal{M}\{\xi_1 \leq \Phi_1^{-1}(\alpha)\} \wedge \mathcal{M}\{\xi_2 \geq \Phi_2^{-1}(1-\alpha)\} \\
&= \alpha \wedge \alpha = \alpha.
\end{aligned}$$

On the other hand, the event $\{\xi \leq \Psi^{-1}(\alpha)\}$ is a subset of either $\mathcal{M}\{\xi_1 \leq \Phi_1^{-1}(\alpha)\}$ or $\{\xi_2 \geq \Phi_2^{-1}(1-\alpha)\}$. Thus $\mathcal{M}\{\xi \leq \Psi^{-1}(\alpha)\} \leq \alpha$. It follows that

$$\mathcal{M}\{\xi \leq \Psi^{-1}(\alpha)\} = \alpha.$$

In other words, Ψ is just the uncertainty distribution of ξ. The theorem is proved.

99 Method 1.5 *Assume $\xi_1, \xi_2, \cdots, \xi_n$ are independent uncertain variables, and each ξ_i is represented on a computer by*

0.01	0.02	0.03	$\cdots$	0.99
x_1^i	x_2^i	x_3^i	$\cdots$	x_{99}^i

(1.134)

If the function $f(x_1, x_2, \cdots, x_n)$ is strictly increasing with respect to $x_1, x_2, \cdots, x_m$ and strictly decreasing with $x_{m+1}, x_{m+2}, \cdots, x_n$, then the uncertain variable $f(\xi_1, \xi_2, \cdots, \xi_n)$ is

0.01	$\cdots$	0.99
$f(x_1^1, \cdots, x_1^m, x_{99}^{m+1}, \cdots, x_{99}^n)$	$\cdots$	$f(x_{99}^1, \cdots, x_{99}^m, x_1^{m+1}, \cdots, x_1^n)$

Example 1.33: Let ξ_1 and ξ_2 be independent uncertain variables with regular uncertainty distributions Φ_1 and Φ_2, respectively. Then the uncertainty distribution Ψ of the difference $\xi_1 - \xi_2$ is

$$\Psi(x) = \sup_{x_1 - x_2 = x} \Phi_1(x_1) \wedge (1 - \Phi_2(x_2)), \quad x \in \Re \qquad (1.135)$$

whose inverse function is

$$\Psi^{-1}(\alpha) = \Phi_1^{-1}(\alpha) - \Phi_2^{-1}(1 - \alpha), \quad 0 < \alpha < 1. \qquad (1.136)$$

If ξ_1 and x_2 are respectively represented by

0.01	0.02	0.03	$\cdots$	0.99
x_1^1	x_2^1	x_3^1	$\cdots$	x_{99}^1

(1.137)

0.01	0.02	0.03	$\cdots$	0.99
x_1^2	x_2^2	x_3^2	$\cdots$	x_{99}^2

(1.138)

then the 99 method yields that $\xi_1 - \xi_2$ is

0.01	0.02	0.03	$\cdots$	0.99
$x_1^1 - x_{99}^2$	$x_2^1 - x_{98}^2$	$x_3^1 - x_{97}^2$	$\cdots$	$x_{99}^1 - x_1^2$

(1.139)

Example 1.34: Let ξ_1 and ξ_2 be independent and positive uncertain variables with regular uncertainty distributions Φ_1 and Φ_2, respectively. Then the uncertainty distribution Ψ of the quotient ξ_1 / ξ_2 is

$$\Psi(x) = \sup_{x_1 / x_2 = x} \Phi_1(x_1) \wedge (1 - \Phi_2(x_2)), \quad x > 0 \qquad (1.140)$$

whose inverse function is

$$\Psi^{-1}(\alpha) = \Phi_1^{-1}(\alpha) / \Phi_2^{-1}(1 - \alpha), \quad 0 < \alpha < 1. \qquad (1.141)$$

If ξ_1 and ξ_2 are respectively represented by

0.01	0.02	0.03	$\cdots$	0.99
x_1^1	x_2^1	x_3^1	$\cdots$	x_{99}^1

(1.142)

0.01	0.02	0.03	$\cdots$	0.99
x_1^2	x_2^2	x_3^2	$\cdots$	x_{99}^2

(1.143)

then the 99 method yields that ξ_1/ξ_2 is

0.01	0.02	0.03	$\cdots$	0.99
x_1^1/x_{99}^2	x_2^1/x_{98}^2	x_3^1/x_{97}^2	$\cdots$	x_{99}^1/x_1^2

(1.144)

Example 1.35: Assume ξ_1, ξ_2, ξ_3 are independent and positive uncertain variables with regular uncertainty distributions Φ_1, Φ_2, Φ_3, respectively. Then the uncertainty distribution of $\xi_1/(\xi_2 + \xi_3)$ is

$$\Psi(x) = \sup_{x_1/(x_2+x_3)=x} \Phi_1(x_1) \wedge (1 - \Phi_2(x_2)) \wedge (1 - \Phi_3(x_3)), \quad x > 0 \quad (1.145)$$

whose inverse function is

$$\Psi^{-1}(\alpha) = \Phi_1^{-1}(\alpha)/(\Phi_2^{-1}(1-\alpha) + \Phi_3^{-1}(1-\alpha)), \quad 0 < \alpha < 1. \quad (1.146)$$

If ξ_1, ξ_2, ξ_3 are respectively represented by

0.01	0.02	0.03	$\cdots$	0.99
x_1^1	x_2^1	x_3^1	$\cdots$	x_{99}^1

(1.147)

0.01	0.02	0.03	$\cdots$	0.99
x_1^2	x_2^2	x_3^2	$\cdots$	x_{99}^2

(1.148)

0.01	0.02	0.03	$\cdots$	0.99
x_1^3	x_2^3	x_3^3	$\cdots$	x_{99}^3

(1.149)

then the 99 method yields that the uncertain variable $\xi_1/(\xi_2 + \xi_3)$ is

0.01	0.02	$\cdots$	0.99
$x_1^1/(x_{99}^2 + x_{99}^3)$	$x_2^1/(x_{98}^2 + x_{98}^3)$	$\cdots$	$x_{99}^1/(x_1^2 + x_1^3)$

(1.150)

1.7   Expected Value

Expected value is the average value of uncertain variable in the sense of uncertain measure, and represents the size of uncertain variable.

Definition 1.21 *(Liu [127]) Let ξ be an uncertain variable. Then the expected value of ξ is defined by*

$$E[\xi] = \int_0^{+\infty} \mathcal{M}\{\xi \geq r\}\mathrm{d}r - \int_{-\infty}^0 \mathcal{M}\{\xi \leq r\}\mathrm{d}r \quad (1.151)$$

provided that at least one of the two integrals is finite.

Theorem 1.32 *Let ξ be an uncertain variable with uncertainty distribution Φ. If the expected value exists, then*

$$E[\xi] = \int_0^{+\infty} (1 - \Phi(x))\mathrm{d}x - \int_{-\infty}^0 \Phi(x)\mathrm{d}x. \qquad (1.152)$$

Proof: It follows from the definitions of expected value operator and uncertainty distribution that

$$E[\xi] = \int_0^{+\infty} \mathcal{M}\{\xi \geq r\}\mathrm{d}r - \int_{-\infty}^0 \mathcal{M}\{\xi \leq r\}\mathrm{d}r$$

$$= \int_0^{+\infty} (1 - \Phi(x))\mathrm{d}x - \int_{-\infty}^0 \Phi(x)\mathrm{d}x.$$

See Figure 1.8. The theorem is proved.
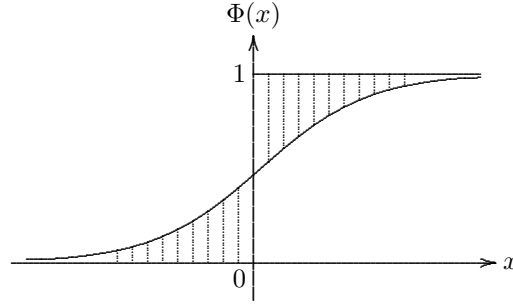
Figure 1.8: Expected Value $E[\xi] = \int_0^{+\infty} (1 - \Phi(x))\mathrm{d}x - \int_{-\infty}^0 \Phi(x)\mathrm{d}x$

Theorem 1.33 *Let ξ be an uncertain variable with regular uncertainty distribution Φ. If the expected value exists, then*

$$E[\xi] = \int_0^1 \Phi^{-1}(\alpha)\mathrm{d}\alpha. \qquad (1.153)$$

Proof: It follows from the definitions of expected value operator and uncertainty distribution that

$$E[\xi] = \int_0^{+\infty} \mathcal{M}\{\xi \geq r\}\mathrm{d}r - \int_{-\infty}^0 \mathcal{M}\{\xi \leq r\}\mathrm{d}r$$

$$= \int_{\Phi(0)}^1 \Phi^{-1}(\alpha)\mathrm{d}\alpha + \int_0^{\Phi(0)} \Phi^{-1}(\alpha)\mathrm{d}\alpha = \int_0^1 \Phi^{-1}(\alpha)\mathrm{d}\alpha.$$

See Figure 1.9. The theorem is proved.
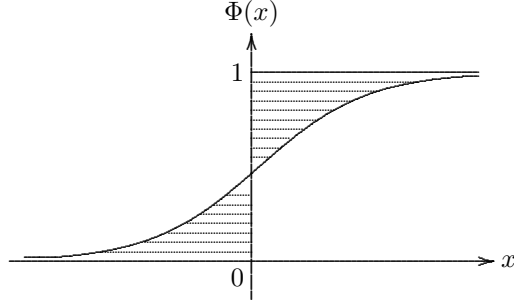
Figure 1.9: Expected Value $E[\xi] = \int_0^1 \Phi^{-1}(\alpha)\mathrm{d}\alpha$

Theorem 1.34 *Let ξ be an uncertain variable with uncertainty distribution Φ. If the expected value exists, then*

$$E[\xi] = \int_{-\infty}^{+\infty} x\mathrm{d}\Phi(x). \qquad (1.154)$$

Proof: Without loss of generality, we suppose that Φ is regular. Otherwise, we may give it a small perturbation such that it becomes regular. It follows from Theorem 1.33 that

$$E[\xi] = \int_0^1 \Phi^{-1}(\alpha)\mathrm{d}\alpha.$$

Now write $\Phi^{-1}(\alpha) = x$. Then we immediately have $\alpha = \Phi(x)$. The change of variable of integral produces (1.154). The theorem is verified.

Example 1.36: Suppose that ξ is a simple uncertain variable taking values in $\{x_1, x_2, \cdots, x_m\}$ with $x_1 < x_2 < \cdots < x_m$. That is, the uncertainty distribution Φ of ξ is a step function jumping only at $x_1, x_2, \cdots, x_m$, i.e.,

$$\Phi(x) = \begin{cases} \Phi(x_0), & \text{if } x < x_1 \\ \Phi(x_1), & \text{if } x_1 \leq x < x_2 \\ \cdots \\ \Phi(x_{m-1}), & \text{if } x_{m-1} \leq x < x_m \\ \Phi(x_m), & \text{if } x \geq x_m. \end{cases}$$

Note that $\Phi(x_0) = 0$ and $\Phi(x_m) = 1$. If $x_1 \geq 0$, then the expected value is

$$
\begin{aligned}
E[\xi] &= \int_0^{x_1} 1 \mathrm{d}x + \sum_{i=1}^{m-1} \int_{x_i}^{x_{i+1}} (1 - \Phi(x_i)) \mathrm{d}x + \int_{x_m}^{+\infty} 0 \mathrm{d}x \\
&= x_1 + \sum_{i=1}^{m-1} (1 - \Phi(x_i))(x_{i+1} - x_i) + 0 \\
&= \sum_{i=1}^{m} (\Phi(x_i) - \Phi(x_{i-1})) x_i.
\end{aligned}
$$

If $x_m \leq 0$, then the expected value is

$$
\begin{aligned}
E[\xi] &= -\int_{-\infty}^{x_1} 0 \mathrm{d}x - \sum_{i=1}^{m-1} \int_{x_i}^{x_{i+1}} \Phi(x_i) \mathrm{d}x - \int_{x_m}^{0} 1 \mathrm{d}x \\
&= 0 - \sum_{i=1}^{m-1} \Phi(x_i)(x_{i+1} - x_i) + x_m \\
&= \sum_{i=1}^{m} (\Phi(x_i) - \Phi(x_{i-1})) x_i.
\end{aligned}
$$

If there exists an index k such that $x_k \leq 0 \leq x_{k+1}$, then the expected value is

$$
\begin{aligned}
E[\xi] &= \int_0^{x_{k+1}} (1 - \Phi(x_k)) \mathrm{d}x + \sum_{i=k+1}^{m-1} \int_{x_i}^{x_{i+1}} (1 - \Phi(x_i)) \mathrm{d}x \\
&\quad - \sum_{i=1}^{k-1} \int_{x_i}^{x_{i+1}} \Phi(x_i) \mathrm{d}x - \int_{x_k}^{0} \Phi(x_k) \mathrm{d}x \\
&= X_{k+1}(1 - \Phi(x_k)) + \sum_{i=k+1}^{m-1} (1 - \Phi(x_i))(x_{i+1} - x_i) \\
&\quad - \sum_{i=1}^{k-1} \Phi(x_i)(x_{i+1} - x_i) + x_k \Phi(x_k) \\
&= \sum_{i=1}^{m} (\Phi(x_i) - \Phi(x_{i-1})) x_i.
\end{aligned}
$$

Thus we always have the expected value

$$
E[\xi] = \sum_{i=1}^{m} (\Phi(x_i) - \Phi(x_{i-1})) x_i \tag{1.155}
$$

where $\Phi(x_0) = 0$ and $\Phi(x_m) = 1$.

Example 1.37: Let $\xi \sim \mathcal{L}(a, b)$ be a linear uncertain variable. If $a \geq 0$, then the expected value is

$$E[\xi] = \left(\int_0^a 1 \mathrm{d}x + \int_a^b \left(1 - \frac{x - a}{b - a} \right) \mathrm{d}x + \int_b^{+\infty} 0 \mathrm{d}x \right) - \int_{-\infty}^0 0 \mathrm{d}x = \frac{a + b}{2}.$$

If $b \leq 0$, then the expected value is

$$E[\xi] = \int_0^{+\infty} 0 \mathrm{d}x - \left(\int_{-\infty}^a 0 \mathrm{d}x + \int_a^b \frac{x - a}{b - a} \mathrm{d}x + \int_b^0 1 \mathrm{d}x \right) = \frac{a + b}{2}.$$

If $a < 0 < b$, then the expected value is

$$E[\xi] = \int_0^b \left(1 - \frac{x - a}{b - a} \right) \mathrm{d}x - \int_a^0 \frac{x - a}{b - a} \mathrm{d}x = \frac{a + b}{2}.$$

Thus we always have the expected value

$$E[\xi] = \frac{a + b}{2}. \tag{1.156}$$

Example 1.38: The zigzag uncertain variable $\xi \sim \mathcal{Z}(a, b, c)$ has an expected value

$$E[\xi] = \frac{a + 2b + c}{4}. \tag{1.157}$$

Example 1.39: The normal uncertain variable $\xi \sim \mathcal{N}(e, \sigma)$ has an expected value e, i.e.,

$$E[\xi] = e. \tag{1.158}$$

Example 1.40: If $\sigma < \pi/\sqrt{3}$, then the lognormal uncertain variable $\xi \sim \mathcal{LOGN}(e, \sigma)$ has an expected value

$$E[\xi] = \sqrt{3}\sigma \exp(e) \csc(\sqrt{3}\sigma). \tag{1.159}$$

Otherwise, $E[\xi] = +\infty$.

Linearity of Expected Value Operator

Theorem 1.35 *Let ξ and η be independent uncertain variables with finite expected values. Then for any real numbers a and b, we have*

$$E[a\xi + b\eta] = aE[\xi] + bE[\eta]. \tag{1.160}$$

Proof: Without loss of generality, suppose that ξ and η are independent uncertain variables with regular uncertainty distributions Φ and Ψ, respectively. Otherwise, we may give them a small perturbation such that the uncertainty distributions are regular.

STEP 1: We first prove $E[a\xi] = aE[\xi]$. If $a = 0$, then the equation holds trivially. If $a > 0$, then the uncertainty distribution Υ of $a\xi$ is determined by

$$\Upsilon^{-1}(\alpha) = a\Phi^{-1}(\alpha).$$

It follows from Theorem 1.33 that

$$E[a\xi] = \int_0^1 a\Phi^{-1}(\alpha)\mathrm{d}\alpha = a\int_0^1 \Phi^{-1}(\alpha)\mathrm{d}\alpha = aE[\xi].$$

If $a < 0$, then the uncertainty distribution Υ of $a\xi$ is determined by

$$\Upsilon^{-1}(\alpha) = a\Phi^{-1}(1 - \alpha).$$

It follows from Theorem 1.33 that

$$E[a\xi] = \int_0^1 a\Phi^{-1}(1 - \alpha)\mathrm{d}\alpha = a\int_0^1 \Phi^{-1}(\alpha)\mathrm{d}\alpha = aE[\xi].$$

Thus we always have $E[a\xi] = aE[\xi]$.

STEP 2: We prove $E[\xi + \eta] = E[\xi] + E[\eta]$. The uncertainty distribution Υ of the sum $\xi + \eta$ is described by

$$\Upsilon^{-1}(\alpha) = \Phi^{-1}(\alpha) + \Psi^{-1}(\alpha).$$

It follows from Theorem 1.33 that

$$E[\xi + \eta] = \int_0^1 \Upsilon^{-1}(\alpha)\mathrm{d}\alpha = \int_0^1 \Phi^{-1}(\alpha)\mathrm{d}\alpha + \int_0^1 \Psi^{-1}(\alpha)\mathrm{d}\alpha = E[\xi] + E[\eta].$$

STEP 3: Finally, for any real numbers a and b, it follows from Steps 1 and 2 that

$$E[a\xi + b\eta] = E[a\xi] + E[b\eta] = aE[\xi] + bE[\eta].$$

The theorem is proved.

Example 1.41: Generally speaking, the expected value operator is not necessarily linear if ξ and η are not independent. For example, take $(\Gamma, \mathcal{L}, \mathcal{M})$ to be $\{\gamma_1, \gamma_2, \gamma_3\}$ with $\mathcal{M}\{\gamma_1\} = 0.7$, $\mathcal{M}\{\gamma_2\} = 0.3$, $\mathcal{M}\{\gamma_3\} = 0.2$, $\mathcal{M}\{\gamma_1, \gamma_2\} = 0.8$, $\mathcal{M}\{\gamma_1, \gamma_3\} = 0.7$, $\mathcal{M}\{\gamma_2, \gamma_3\} = 0.3$. The uncertain variables are defined by

$$\xi_1(\gamma) = \begin{cases} 1, & \text{if } \gamma = \gamma_1 \\ 0, & \text{if } \gamma = \gamma_2 \\ 2, & \text{if } \gamma = \gamma_3, \end{cases} \qquad \xi_2(\gamma) = \begin{cases} 0, & \text{if } \gamma = \gamma_1 \\ 2, & \text{if } \gamma = \gamma_2 \\ 3, & \text{if } \gamma = \gamma_3. \end{cases}$$

Note that ξ_1 and ξ_2 are not independent, and their sum is

$$(\xi_1 + \xi_2)(\gamma) = \begin{cases} 1, & \text{if } \gamma = \gamma_1 \\ 2, & \text{if } \gamma = \gamma_2 \\ 5, & \text{if } \gamma = \gamma_3. \end{cases}$$

Thus $E[\xi_1] = 0.9$, $E[\xi_2] = 0.8$, and $E[\xi_1 + \xi_2] = 1.9$. This fact implies that

$$E[\xi_1 + \xi_2] > E[\xi_1] + E[\xi_2].$$

If the uncertain variables are defined by

$$\eta_1(\gamma) = \begin{cases} 0, & \text{if } \gamma = \gamma_1 \\ 1, & \text{if } \gamma = \gamma_2 \\ 2, & \text{if } \gamma = \gamma_3, \end{cases} \qquad \eta_2(\gamma) = \begin{cases} 0, & \text{if } \gamma = \gamma_1 \\ 3, & \text{if } \gamma = \gamma_2 \\ 1, & \text{if } \gamma = \gamma_3. \end{cases}$$

Then we have

$$(\eta_1 + \eta_2)(\gamma) = \begin{cases} 0, & \text{if } \gamma = \gamma_1 \\ 4, & \text{if } \gamma = \gamma_2 \\ 3, & \text{if } \gamma = \gamma_3. \end{cases}$$

Thus $E[\eta_1] = 0.5$, $E[\eta_2] = 0.9$, and $E[\eta_1 + \eta_2] = 1.2$. This fact implies that

$$E[\eta_1 + \eta_2] < E[\eta_1] + E[\eta_2].$$

Expected Value of Function of Single Uncertain Variable

Let ξ be an uncertain variable, and $f : \Re \to \Re$ a function. Then the expected value of $f(\xi)$ is

$$E[f(\xi)] = \int_0^{+\infty} \mathcal{M}\{f(\xi) \geq r\}\mathrm{d}r - \int_{-\infty}^0 \mathcal{M}\{f(\xi) \leq r\}\mathrm{d}r.$$

For random case, it has been proved that the expected value $E[f(\xi)]$ is the Lebesgue-Stieltjes integral of $f(x)$ with respect to the uncertainty distribution Φ of ξ if the integral exists. However, generally speaking, it is not true for uncertain case.

Example 1.42: We consider an uncertain variable ξ whose first identification function is given by

$$\lambda(x) = \begin{cases} 0.3, & \text{if } -1 \leq x < 0 \\ 0.5, & \text{if } 0 \leq x \leq 1. \end{cases}$$

Then the expected value $E[\xi^2] = 0.5$. However, the uncertainty distribution of ξ is

$$\Phi(x) = \begin{cases} 0, & \text{if } x < -1 \\ 0.3, & \text{if } -1 \leq x < 0 \\ 0.5, & \text{if } 0 \leq x < 1 \\ 1, & \text{if } x \geq 1 \end{cases}$$

and the Lebesgue-Stieltjes integral

$$\int_{-\infty}^{+\infty} x^2 \mathrm{d}\Phi(x) = (-1)^2 \times 0.3 + 0^2 \times 0.2 + 1^2 \times 0.5 = 0.8 \neq E[\xi^2].$$

Theorem 1.36 *(Liu and Ha [138]) Let ξ be an uncertain variable whose uncertainty distribution Φ. If $f(x)$ is a strictly monotone function such that the expected value $E[f(\xi)]$ exists, then*

$$E[f(\xi)] = \int_{-\infty}^{+\infty} f(x)\mathrm{d}\Phi(x). \tag{1.161}$$

Proof: We first suppose that $f(x)$ is a strictly increasing function. Then $f(\xi)$ has an uncertainty distribution $\Phi(f^{-1}(x))$. It follows from the change of variable of integral that

$$E[f(\xi)] = \int_{-\infty}^{+\infty} x\mathrm{d}\Phi(f^{-1}(x)) = \int_{-\infty}^{+\infty} f(x)\mathrm{d}\Phi(x).$$

If $f(x)$ is a strictly decreasing function, then $-f(x)$ is a strictly increasing function. Hence

$$E[f(\xi)] = -E[-f(\xi)] = -\int_{-\infty}^{+\infty} -f(x)\mathrm{d}\Phi(x) = \int_{-\infty}^{+\infty} f(y)\mathrm{d}\Phi(y).$$

The theorem is verified.

Example 1.43: Let ξ be a positive linear uncertain variable $\mathcal{L}(a, b)$. Then its uncertainty distribution is $\Phi(x) = (x - a)/(b - a)$. Thus

$$E[\xi^2] = \int_a^b x^2 \mathrm{d}\Phi(x) = \frac{a^2 + b^2 + ab}{3}.$$

Example 1.44: Let ξ be a positive linear uncertain variable $\mathcal{L}(a, b)$. Then its uncertainty distribution is $\Phi(x) = (x - a)/(b - a)$. Thus

$$E[\exp(\xi)] = \int_a^b \exp(x)\mathrm{d}\Phi(x) = \frac{\exp(b) - \exp(a)}{b - a}.$$

Theorem 1.37 *(Liu and Ha [138]) Assume ξ is an uncertain variable with regular uncertainty distribution Φ. If $f(x)$ is a strictly monotone function such that the expected value $E[f(\xi)]$ exists, then*

$$E[f(\xi)] = \int_0^1 f(\Phi^{-1}(\alpha))\mathrm{d}\alpha. \tag{1.162}$$

Proof: Suppose that f is a strictly increasing function. It follows from Theorem 1.25 that the uncertainty distribution of $f(\xi)$ is described by

$$\Psi^{-1}(\alpha) = f(\Phi^{-1}(\alpha)).$$

By using Theorem 1.33, the equation (1.162) is proved. When f is a strictly decreasing function, it follows from Theorem 1.30 that the uncertainty distribution of $f(\xi)$ is described by

$$\Psi^{-1}(\alpha) = f(\Phi^{-1}(1 - \alpha)).$$

By using Theorem 1.33 and the change of variable of integral, we get the equation (1.162). The theorem is verified.

Example 1.45: Let ξ be a nonnegative uncertain variable with regular uncertainty distribution Φ. Then

$$E[\sqrt{\xi}] = \int_0^1 \sqrt{\Phi^{-1}(\alpha)}\mathrm{d}\alpha. \tag{1.163}$$

Example 1.46: Let ξ be a positive uncertain variable with regular uncertainty distribution Φ. Then

$$E\left[\frac{1}{\xi}\right] = \int_0^1 \frac{1}{\Phi^{-1}(1 - \alpha)}\mathrm{d}\alpha = \int_0^1 \frac{1}{\Phi^{-1}(\alpha)}\mathrm{d}\alpha. \tag{1.164}$$

Expected Value of Function of Multiple Uncertain Variables

Theorem 1.38 *(Liu and Ha [138]) Assume $\xi_1, \xi_2, \cdots, \xi_n$ are independent uncertain variables with regular uncertainty distributions $\Phi_1, \Phi_2, \cdots, \Phi_n$, respectively. If $f : \Re^n \to \Re$ is a strictly monotone function, then the uncertain variable $\xi = f(\xi_1, \xi_2, \cdots, \xi_n)$ has an expected value*

$$E[\xi] = \int_0^1 f(\Phi_1^{-1}(\alpha), \Phi_2^{-1}(\alpha), \cdots, \Phi_n^{-1}(\alpha))\mathrm{d}\alpha \tag{1.165}$$

provided that the expected value $E[\xi]$ exists.

Proof: Suppose that f is a strictly increasing function. It follows from Theorem 1.25 that the uncertainty distribution of ξ has an inverse function

$$\Psi^{-1}(\alpha) = f(\Phi_1^{-1}(\alpha), \Phi_2^{-1}(\alpha), \cdots, \Phi_n^{-1}(\alpha)).$$

By using Theorem 1.33, we obtain (1.165). When f is a strictly decreasing function, it follows from Theorem 1.30 that the uncertainty distribution of ξ has an inverse function

$$\Psi^{-1}(\alpha) = f(\Phi_1^{-1}(1 - \alpha), \Phi_2^{-1}(1 - \alpha), \cdots, \Phi_n^{-1}(1 - \alpha)).$$

By using Theorem 1.33 and the change of variable of integral, we obtain (1.165). The theorem is proved.

Example 1.47: Let ξ and η be independent and nonnegative uncertain variables with regular uncertainty distributions Φ and Ψ, respectively. Then

$$E[\xi\eta] = \int_0^1 \Phi^{-1}(\alpha)\Psi^{-1}(\alpha)\mathrm{d}\alpha. \tag{1.166}$$

Exercise 1.1. What is the expected value of an alternating monotone function of uncertain variables?

Exercise 1.2. Let ξ and η be independent and positive uncertain variables with regular uncertainty distributions Φ and Ψ, respectively. Prove

$$E\left[\frac{\xi}{\eta}\right] = \int_0^1 \frac{\Phi^{-1}(\alpha)}{\Psi^{-1}(1-\alpha)}\mathrm{d}\alpha. \tag{1.167}$$

1.8 Variance

The variance of uncertain variable provides a measure of the spread of the distribution around its expected value. A small value of variance indicates that the uncertain variable is tightly concentrated around its expected value; and a large value of variance indicates that the uncertain variable has a wide spread around its expected value.

Definition 1.22 *(Liu [127]) Let ξ be an uncertain variable with finite expected value e. Then the variance of ξ is defined by $V[\xi] = E[(\xi - e)^2]$.*

Let ξ be an uncertain variable with expected value e. If we only know its uncertainty distribution Φ, then the variance

$$
\begin{aligned}
V[\xi] &= \int_0^{+\infty} \mathcal{M}\{(\xi - e)^2 \geq r\}\mathrm{d}r \\
&= \int_0^{+\infty} \mathcal{M}\{(\xi \geq e + \sqrt{r}) \cup (\xi \leq e - \sqrt{r})\}\mathrm{d}r \\
&\leq \int_0^{+\infty} (\mathcal{M}\{\xi \geq e + \sqrt{r}\} + \mathcal{M}\{\xi \leq e - \sqrt{r}\})\mathrm{d}r \\
&= \int_0^{+\infty} (1 - \Phi(e + \sqrt{r}) + \Phi(e - \sqrt{r}))\mathrm{d}r \\
&= \int_e^{+\infty} 2(r - e)(1 - \Phi(r) + \Phi(2e - r))\mathrm{d}r.
\end{aligned}
$$

For this case, we will stipulate that the variance is

$$V[\xi] = 2 \int_e^{+\infty} (r - e)(1 - \Phi(r) + \Phi(2e - r)) \mathrm{d}r. \qquad (1.168)$$

Mention that this is a stipulation rather than a precise formula!

Example 1.48: It has been verified that the linear uncertain variable $\xi \sim \mathcal{L}(a, b)$ has an expected value $(a+b)/2$. Note that the uncertainty distribution is $\Phi(x) = (x - a)/(b - a)$ when $a \leq x \leq b$. It follows from the stipulation (1.168) that the variance is

$$V[\xi] = 2 \int_{(a+b)/2}^b \left(r - \frac{a + b}{2} \right) \left(1 - \frac{r - a}{b - a} + \frac{b - r}{b - a} \right) \mathrm{d}r = \frac{(b - a)^2}{12}.$$

In fact, a precise conclusion is $(b - a)^2/24 \leq V[\xi] \leq (b - a)^2/12$.

Example 1.49: It has been verified that the normal uncertain variable $\xi \sim \mathcal{N}(e, \sigma)$ has expected value e. It follows from the stipulation (1.168) that the variance is

$$V[\xi] = \sigma^2. \qquad (1.169)$$

In fact, a precise conclusion is $\sigma^2/2 \leq V[\xi] \leq \sigma^2$.

Theorem 1.39 *If ξ is an uncertain variable with finite expected value, a and b are real numbers, then $V[a\xi + b] = a^2 V[\xi]$.*

Proof: It follows from the definition of variance that

$$V[a\xi + b] = E\left[(a\xi + b - aE[\xi] - b)^2 \right] = a^2 E[(\xi - E[\xi])^2] = a^2 V[\xi].$$

Theorem 1.40 *Let ξ be an uncertain variable with expected value e. Then $V[\xi] = 0$ if and only if $\mathcal{M}\{\xi = e\} = 1$.*

Proof: If $V[\xi] = 0$, then $E[(\xi - e)^2] = 0$. Note that

$$E[(\xi - e)^2] = \int_0^{+\infty} \mathcal{M}\{(\xi - e)^2 \geq r\} \mathrm{d}r$$

which implies $\mathcal{M}\{(\xi - e)^2 \geq r\} = 0$ for any $r > 0$. Hence we have

$$\mathcal{M}\{(\xi - e)^2 = 0\} = 1.$$

That is, $\mathcal{M}\{\xi = e\} = 1$. Conversely, if $\mathcal{M}\{\xi = e\} = 1$, then we have $\mathcal{M}\{(\xi - e)^2 = 0\} = 1$ and $\mathcal{M}\{(\xi - e)^2 \geq r\} = 0$ for any $r > 0$. Thus

$$V[\xi] = \int_0^{+\infty} \mathcal{M}\{(\xi - e)^2 \geq r\} \mathrm{d}r = 0.$$

The theorem is proved.

Maximum Variance Theorem

Let ξ be an uncertain variable that takes values in $[a, b]$, but whose uncertainty distribution is otherwise arbitrary. If its expected value is given, what is the possible maximum variance? The maximum variance theorem will answer this question, thus playing an important role in treating games against nature.

Theorem 1.41 *Let f be a convex function on $[a, b]$, and ξ an uncertain variable that takes values in $[a, b]$ and has expected value e. Then*

$$E[f(\xi)] \leq \frac{b-e}{b-a}f(a) + \frac{e-a}{b-a}f(b). \tag{1.170}$$

Proof: For each $\gamma \in \Gamma$, we have $a \leq \xi(\gamma) \leq b$ and

$$\xi(\gamma) = \frac{b-\xi(\gamma)}{b-a}a + \frac{\xi(\gamma)-a}{b-a}b.$$

It follows from the convexity of f that

$$f(\xi(\gamma)) \leq \frac{b-\xi(\gamma)}{b-a}f(a) + \frac{\xi(\gamma)-a}{b-a}f(b).$$

Taking expected values on both sides, we obtain the inequality.

Theorem 1.42 *(Maximum Variance Theorem) Let ξ be an uncertain variable that takes values in $[a, b]$ and has expected value e. Then*

$$V[\xi] \leq (e-a)(b-e) \tag{1.171}$$

and equality holds if the uncertain variable ξ is determined by

$$\mathcal{M}\{\xi = x\} = \begin{cases} \dfrac{b-e}{b-a}, & \text{if } x = a \\[2mm] \dfrac{e-a}{b-a}, & \text{if } x = b. \end{cases} \tag{1.172}$$

Proof: It follows from Theorem 1.41 immediately by defining $f(x) = (x-e)^2$. It is also easy to verify that the uncertain variable determined by (1.172) has variance $(e-a)(b-e)$. The theorem is proved.

1.9 Moments

Definition 1.23 *(Liu [127]) Let ξ be an uncertain variable. Then for any positive integer k,*
(a) the expected value $E[\xi^k]$ is called the kth moment;
(b) the expected value $E[|\xi|^k]$ is called the kth absolute moment;
(c) the expected value $E[(\xi - E[\xi])^k]$ is called the kth central moment;
(d) the expected value $E[|\xi - E[\xi]|^k]$ is called the kth absolute central moment.

Note that the first central moment is always 0, the first moment is just the expected value, and the second central moment is just the variance.

Theorem 1.43 *Let ξ be a nonnegative uncertain variable, and k a positive number. Then the k-th moment*

$$E[\xi^k] = k \int_0^{+\infty} r^{k-1} \mathcal{M}\{\xi \geq r\} \mathrm{d}r. \tag{1.173}$$

Proof: It follows from the nonnegativity of ξ that

$$E[\xi^k] = \int_0^\infty \mathcal{M}\{\xi^k \geq x\} \mathrm{d}x = \int_0^\infty \mathcal{M}\{\xi \geq r\} \mathrm{d}r^k = k \int_0^\infty r^{k-1} \mathcal{M}\{\xi \geq r\} \mathrm{d}r.$$

The theorem is proved.

Theorem 1.44 *Let ξ be an uncertain variable, and t a positive number. If $E[|\xi|^t] < \infty$, then*

$$\lim_{x \to \infty} x^t \mathcal{M}\{|\xi| \geq x\} = 0. \tag{1.174}$$

Conversely, if (1.174) holds for some positive number t, then $E[|\xi|^s] < \infty$ for any $0 \leq s < t$.

Proof: It follows from the definition of expected value operator that

$$E[|\xi|^t] = \int_0^{+\infty} \mathcal{M}\{|\xi|^t \geq r\} \mathrm{d}r < \infty.$$

Thus we have

$$\lim_{x \to \infty} \int_{x^t/2}^{+\infty} \mathcal{M}\{|\xi|^t \geq r\} \mathrm{d}r = 0.$$

The equation (1.174) is proved by the following relation,

$$\int_{x^t/2}^{+\infty} \mathcal{M}\{|\xi|^t \geq r\} \mathrm{d}r \geq \int_{x^t/2}^{x^t} \mathcal{M}\{|\xi|^t \geq r\} \mathrm{d}r \geq \frac{1}{2} x^t \mathcal{M}\{|\xi| \geq x\}.$$

Conversely, if (1.174) holds, then there exists a number $a > 0$ such that

$$x^t \mathcal{M}\{|\xi| \geq x\} \leq 1, \ \forall x \geq a.$$

Thus we have

$$
\begin{aligned}
E[|\xi|^s] &= \int_0^a \mathcal{M}\{|\xi|^s \geq r\} \mathrm{d}r + \int_a^{+\infty} \mathcal{M}\{|\xi|^s \geq r\} \mathrm{d}r \\
&= \int_0^a \mathcal{M}\{|\xi|^s \geq r\} \mathrm{d}r + \int_a^{+\infty} sr^{s-1} \mathcal{M}\{|\xi| \geq r\} \mathrm{d}r \\
&\leq \int_0^a \mathcal{M}\{|\xi|^s \geq r\} \mathrm{d}r + s \int_a^{+\infty} r^{s-t-1} \mathrm{d}r \\
&< +\infty. \quad \left(\text{by } \int_a^{+\infty} r^p \mathrm{d}r < \infty \text{ for any } p < -1 \right)
\end{aligned}
$$

The theorem is proved.

Theorem 1.45 *Let ξ be an uncertain variable that takes values in $[a,b]$ and has expected value e. Then for any positive integer k, the kth absolute moment and kth absolute central moment satisfy the following inequalities,*

$$E[|\xi|^k] \leq \frac{b-e}{b-a}|a|^k + \frac{e-a}{b-a}|b|^k, \tag{1.175}$$

$$E[|\xi - e|^k] \leq \frac{b-e}{b-a}(e-a)^k + \frac{e-a}{b-a}(b-e)^k. \tag{1.176}$$

Proof: It follows from Theorem 1.41 immediately by defining $f(x) = |x|^k$ and $f(x) = |x - e|^k$.

1.10 Critical Values

In order to rank uncertain variables, we may use two critical values: optimistic value and pessimistic value.

Definition 1.24 *(Liu [127]) Let ξ be an uncertain variable, and $\alpha \in (0,1]$. Then*

$$\xi_{\sup}(\alpha) = \sup\left\{r \mid \mathcal{M}\{\xi \geq r\} \geq \alpha\right\} \tag{1.177}$$

is called the α-optimistic value to ξ, and

$$\xi_{\inf}(\alpha) = \inf\left\{r \mid \mathcal{M}\{\xi \leq r\} \geq \alpha\right\} \tag{1.178}$$

is called the α-pessimistic value to ξ.

This means that the uncertain variable ξ will reach upwards of the α-optimistic value $\xi_{\sup}(\alpha)$ with uncertain measure α, and will be below the α-pessimistic value $\xi_{\inf}(\alpha)$ with uncertain measure α.

Theorem 1.46 *Let ξ be an uncertain variable with regular uncertainty distribution Φ. Then its α-optimistic value and α-pessimistic value are*

$$\xi_{\sup}(\alpha) = \Phi^{-1}(1 - \alpha), \tag{1.179}$$

$$\xi_{\inf}(\alpha) = \Phi^{-1}(\alpha). \tag{1.180}$$

Proof: It follows from the definition of α-optimistic value and α-pessimistic value immediately.

Example 1.50: Let ξ be a linear uncertain variable $\mathcal{L}(a,b)$. Then its α-optimistic and α-pessimistic values are

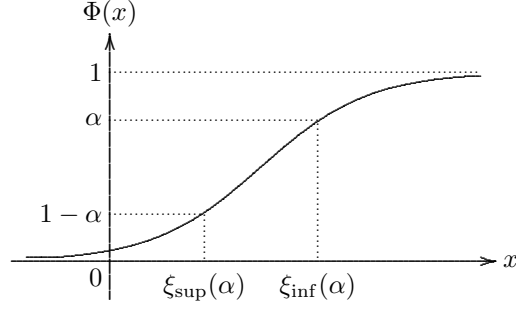$$\xi_{\sup}(\alpha) = \alpha a + (1 - \alpha)b,$$

Figure 1.10: Optimistic Value and Pessimistic Value

$$\xi_{\text{inf}}(\alpha) = (1-\alpha)a + \alpha b.$$

Example 1.51: Let ξ be a zigzag uncertain variable $\mathcal{Z}(a,b,c)$. Then its α-optimistic and α-pessimistic values are

$$\xi_{\text{sup}}(\alpha) = \begin{cases} 2\alpha b + (1-2\alpha)c, & \text{if } \alpha < 0.5 \\ (2\alpha-1)a + (2-2\alpha)b, & \text{if } \alpha \geq 0.5, \end{cases}$$

$$\xi_{\text{inf}}(\alpha) = \begin{cases} (1-2\alpha)a + 2\alpha b, & \text{if } \alpha < 0.5 \\ (2-2\alpha)b + (2\alpha-1)c, & \text{if } \alpha \geq 0.5. \end{cases}$$

Example 1.52: Let ξ be a normal uncertain variable $\mathcal{N}(e,\sigma)$. Then its α-optimistic and α-pessimistic values are

$$\xi_{\text{sup}}(\alpha) = e - \frac{\sigma\sqrt{3}}{\pi}\ln\frac{\alpha}{1-\alpha},$$

$$\xi_{\text{inf}}(\alpha) = e + \frac{\sigma\sqrt{3}}{\pi}\ln\frac{\alpha}{1-\alpha}.$$

Example 1.53: Let ξ be a lognormal uncertain variable $\mathcal{LOGN}(e,\sigma)$. Then its α-optimistic and α-pessimistic values are

$$\xi_{\text{sup}}(\alpha) = \exp(e)\left(\frac{1-\alpha}{\alpha}\right)^{\sqrt{3}\sigma/\pi},$$

$$\xi_{\text{inf}}(\alpha) = \exp(e)\left(\frac{\alpha}{1-\alpha}\right)^{\sqrt{3}\sigma/\pi}.$$

Theorem 1.47 *Let ξ be an uncertain variable, and $\alpha \in (0,1]$. Then for any positive number ε, we have*

$$\mathcal{M}\{\xi \leq \xi_{\text{inf}}(\alpha) + \varepsilon\} \geq \alpha, \quad \mathcal{M}\{\xi \geq \xi_{\text{sup}}(\alpha) - \varepsilon\} \geq \alpha. \tag{1.181}$$

Proof: It follows from the definition of α-pessimistic value that there exists a decreasing sequence $\{x_i\}$ such that $\mathcal{M}\{\xi \leq x_i\} \geq \alpha$ and $x_i \downarrow \xi_{\text{inf}}(\alpha)$ as $i \to \infty$. Thus for any positive number ε, there exists an index i such that $x_i < \xi_{\text{inf}}(\alpha) + \varepsilon$. Hence

$$\mathcal{M}\{\xi \leq \xi_{\text{inf}}(\alpha) + \varepsilon\} \geq \mathcal{M}\{\xi \leq x_i\} \geq \alpha.$$

Similarly, there exists an increasing sequence $\{x_i\}$ such that $\mathcal{M}\{\xi \geq x_i\} \geq \alpha$ and $x_i \uparrow \xi_{\text{sup}}(\alpha)$ as $i \to \infty$. Thus for any positive number ε, there exists an index i such that $x_i > \xi_{\text{sup}}(\alpha) - \varepsilon$. Hence

$$\mathcal{M}\{\xi \geq \xi_{\text{sup}}(\alpha) - \varepsilon\} \geq \mathcal{M}\{\xi \geq x_i\} \geq \alpha.$$

The theorem is proved.

Theorem 1.48 *Let ξ be an uncertain variable, and $\alpha \in (0,1]$. Then we have*
(a) $\xi_{\text{inf}}(\alpha)$ is an increasing and left-continuous function of α;
(b) $\xi_{\text{sup}}(\alpha)$ is a decreasing and left-continuous function of α.

Proof: (a) Let α_1 and α_2 be two numbers with $0 < \alpha_1 < \alpha_2 \leq 1$. Then for any number $r < \xi_{\text{sup}}(\alpha_2)$, we have

$$\mathcal{M}\{\xi \geq r\} \geq \alpha_2 > \alpha_1.$$

Thus, by the definition of optimistic value, we obtain $\xi_{\text{sup}}(\alpha_1) \geq \xi_{\text{sup}}(\alpha_2)$. That is, the value $\xi_{\text{sup}}(\alpha)$ is a decreasing function of α. Next, we prove the left-continuity of $\xi_{\text{inf}}(\alpha)$ with respect to α. Let $\{\alpha_i\}$ be an arbitrary sequence of positive numbers such that $\alpha_i \uparrow \alpha$. Then $\{\xi_{\text{inf}}(\alpha_i)\}$ is an increasing sequence. If the limitation is equal to $\xi_{\text{inf}}(\alpha)$, then the left-continuity is proved. Otherwise, there exists a number z^* such that

$$\lim_{i \to \infty} \xi_{\text{inf}}(\alpha_i) < z^* < \xi_{\text{inf}}(\alpha).$$

Thus $\mathcal{M}\{\xi \leq z^*\} \geq \alpha_i$ for each i. Letting $i \to \infty$, we get $\mathcal{M}\{\xi \leq z^*\} \geq \alpha$. Hence $z^* \geq \xi_{\text{inf}}(\alpha)$. A contradiction proves the left-continuity of $\xi_{\text{inf}}(\alpha)$ with respect to α. The part (b) may be proved similarly.

Theorem 1.49 *Let ξ be an uncertain variable, and $\alpha \in (0,1]$. Then we have*
(a) if $\alpha > 0.5$, then $\xi_{\text{inf}}(\alpha) \geq \xi_{\text{sup}}(\alpha)$;
(b) if $\alpha \leq 0.5$, then $\xi_{\text{inf}}(\alpha) \leq \xi_{\text{sup}}(\alpha)$.

Proof: Part (a): Write $\bar{\xi}(\alpha) = (\xi_{\text{inf}}(\alpha) + \xi_{\text{sup}}(\alpha))/2$. If $\xi_{\text{inf}}(\alpha) < \xi_{\text{sup}}(\alpha)$, then we have

$$1 \geq \mathcal{M}\{\xi < \bar{\xi}(\alpha)\} + \mathcal{M}\{\xi > \bar{\xi}(\alpha)\} \geq \alpha + \alpha > 1.$$

A contradiction proves $\xi_{\text{inf}}(\alpha) \geq \xi_{\text{sup}}(\alpha)$. Part (b): Assume that $\xi_{\text{inf}}(\alpha) > \xi_{\text{sup}}(\alpha)$. It follows from the definition of $\xi_{\text{inf}}(\alpha)$ that $\mathcal{M}\{\xi \leq \bar{\xi}(\alpha)\} < \alpha$.

Similarly, it follows from the definition of $\xi_{\text{sup}}(\alpha)$ that $\mathcal{M}\{\xi \geq \bar{\xi}(\alpha)\} < \alpha$. Thus

$$1 \leq \mathcal{M}\{\xi \leq \bar{\xi}(\alpha)\} + \mathcal{M}\{\xi \geq \bar{\xi}(\alpha)\} < \alpha + \alpha \leq 1.$$

A contradiction proves $\xi_{\text{inf}}(\alpha) \leq \xi_{\text{sup}}(\alpha)$. The theorem is verified.

Theorem 1.50 *Let ξ be an uncertain variable, and $\alpha \in (0, 1]$. Then we have*
(a) if $c \geq 0$, then $(c\xi)_{\text{sup}}(\alpha) = c\xi_{\text{sup}}(\alpha)$ and $(c\xi)_{\text{inf}}(\alpha) = c\xi_{\text{inf}}(\alpha)$;
(b) if $c < 0$, then $(c\xi)_{\text{sup}}(\alpha) = c\xi_{\text{inf}}(\alpha)$ and $(c\xi)_{\text{inf}}(\alpha) = c\xi_{\text{sup}}(\alpha)$.

Proof: (a) If $c = 0$, then the part (a) is obvious. In the case of $c > 0$, we have

$$\begin{aligned}
(c\xi)_{\text{sup}}(\alpha) &= \sup\{r \mid \mathcal{M}\{c\xi \geq r\} \geq \alpha\} \\
&= c\sup\{r/c \mid \mathcal{M}\{\xi \geq r/c\} \geq \alpha\} \\
&= c\xi_{\text{sup}}(\alpha).
\end{aligned}$$

A similar way may prove $(c\xi)_{\text{inf}}(\alpha) = c\xi_{\text{inf}}(\alpha)$. In order to prove the part (b), it suffices to prove that $(-\xi)_{\text{sup}}(\alpha) = -\xi_{\text{inf}}(\alpha)$ and $(-\xi)_{\text{inf}}(\alpha) = -\xi_{\text{sup}}(\alpha)$. In fact, we have

$$\begin{aligned}
(-\xi)_{\text{sup}}(\alpha) &= \sup\{r \mid \mathcal{M}\{-\xi \geq r\} \geq \alpha\} \\
&= -\inf\{-r \mid \mathcal{M}\{\xi \leq -r\} \geq \alpha\} \\
&= -\xi_{\text{inf}}(\alpha).
\end{aligned}$$

Similarly, we may prove that $(-\xi)_{\text{inf}}(\alpha) = -\xi_{\text{sup}}(\alpha)$. The theorem is proved.

Theorem 1.51 *Suppose that ξ and η are independent uncertain variables, and $\alpha \in (0, 1]$. Then we have*

$$(\xi + \eta)_{\text{sup}}(\alpha) = \xi_{\text{sup}}(\alpha) + \eta_{\text{sup}}(\alpha), \quad (\xi + \eta)_{\text{inf}}(\alpha) = \xi_{\text{inf}}(\alpha) + \eta_{\text{inf}}(\alpha),$$

$$(\xi \vee \eta)_{\text{sup}}(\alpha) = \xi_{\text{sup}}(\alpha) \vee \eta_{\text{sup}}(\alpha), \quad (\xi \vee \eta)_{\text{inf}}(\alpha) = \xi_{\text{inf}}(\alpha) \vee \eta_{\text{inf}}(\alpha),$$

$$(\xi \wedge \eta)_{\text{sup}}(\alpha) = \xi_{\text{sup}}(\alpha) \wedge \eta_{\text{sup}}(\alpha), \quad (\xi \wedge \eta)_{\text{inf}}(\alpha) = \xi_{\text{inf}}(\alpha) \wedge \eta_{\text{inf}}(\alpha).$$

Proof: For any given number $\varepsilon > 0$, since ξ and η are independent uncertain variables, we have

$$\begin{aligned}
&\mathcal{M}\{\xi + \eta \geq \xi_{\text{sup}}(\alpha) + \eta_{\text{sup}}(\alpha) - \varepsilon\} \\
\geq\ &\mathcal{M}\left\{(\xi \geq \xi_{\text{sup}}(\alpha) - \varepsilon/2) \cap (\eta \geq \eta_{\text{sup}}(\alpha) - \varepsilon/2)\right\} \\
=\ &\mathcal{M}\{\xi \geq \xi_{\text{sup}}(\alpha) - \varepsilon/2\} \wedge \mathcal{M}\{\eta \geq \eta_{\text{sup}}(\alpha) - \varepsilon/2\} \geq \alpha
\end{aligned}$$

which implies

$$(\xi + \eta)_{\text{sup}}(\alpha) \geq \xi_{\text{sup}}(\alpha) + \eta_{\text{sup}}(\alpha) - \varepsilon. \tag{1.182}$$

On the other hand, by the independence, we have

$$\mathcal{M}\{\xi + \eta \geq \xi_{\sup}(\alpha) + \eta_{\sup}(\alpha) + \varepsilon\}$$
$$\leq \mathcal{M}\left\{(\xi \geq \xi_{\sup}(\alpha) + \varepsilon/2) \cup (\eta \geq \eta_{\sup}(\alpha) + \varepsilon/2)\right\}$$
$$= \mathcal{M}\{\xi \geq \xi_{\sup}(\alpha) + \varepsilon/2\} \vee \mathcal{M}\{\eta \geq \eta_{\sup}(\alpha) + \varepsilon/2\} < \alpha$$

which implies

$$(\xi + \eta)_{\sup}(\alpha) \leq \xi_{\sup}(\alpha) + \eta_{\sup}(\alpha) + \varepsilon. \tag{1.183}$$

It follows from (1.182) and (1.183) that

$$\xi_{\sup}(\alpha) + \eta_{\sup}(\alpha) + \varepsilon \geq (\xi + \eta)_{\sup}(\alpha) \geq \xi_{\sup}(\alpha) + \eta_{\sup}(\alpha) - \varepsilon.$$

Letting $\varepsilon \to 0$, we obtain $(\xi + \eta)_{\sup}(\alpha) = \xi_{\sup}(\alpha) + \eta_{\sup}(\alpha)$. The other equalities may be proved similarly.

Example 1.54: The independence condition cannot be removed in Theorem 1.51. For example, let $\Gamma = \{\gamma_1, \gamma_2\}$, $\mathcal{M}\{\gamma_1\} = \mathcal{M}\{\gamma_2\} = 1/2$, and let uncertain variables ξ and η be defined as

$$\xi(\gamma) = \begin{cases} 0, & \text{if } \gamma = \gamma_1 \\ 1, & \text{if } \gamma = \gamma_2, \end{cases} \qquad \eta(\gamma) = \begin{cases} 1, & \text{if } \gamma = \gamma_1 \\ 0, & \text{if } \gamma = \gamma_2. \end{cases}$$

However, $(\xi + \eta)_{\sup}(0.6) = 1 \neq 0 = \xi_{\sup}(0.6) + \eta_{\sup}(0.6)$.

Theorem 1.52 *(Zuo [253]) Let $\xi_1, \xi_2, \cdots, \xi_n$ be independent uncertain variables with regular uncertainty distributions. If $f : \Re^n \to \Re$ is a continuous and strictly increasing function, then $\xi = f(\xi_1, \xi_2, \cdots, \xi_n)$ is an uncertain variable, and*

$$\xi_{\sup}(\alpha) = f(\xi_{1\,\sup}(\alpha), \xi_{2\,\sup}(\alpha), \cdots, \xi_{n\,\sup}(\alpha)), \tag{1.184}$$

$$\xi_{\inf}(\alpha) = f(\xi_{1\,\inf}(\alpha), \xi_{2\,\inf}(\alpha), \cdots, \xi_{n\,\inf}(\alpha)). \tag{1.185}$$

Proof: Since f is a strictly increasing function, it follows from Theorem 1.25 that the uncertainty distribution Ψ of ξ is determined by

$$\Psi^{-1}(\alpha) = f(\Phi_1^{-1}(\alpha), \Phi_2^{-1}(\alpha), \cdots, \Phi_n^{-1}(\alpha))$$

where $\Phi_1, \Phi_2, \cdots, \Phi_n$ are uncertainty distributions of $\xi_1, \xi_2, \cdots, \xi_n$, respectively. By using Theorem 1.46, we get (1.184) and (1.185). The theorem is proved.

Example 1.55: Let ξ and η be independent and positive uncertain variables with regular uncertainty distributions. Since $f(x, y) = xy$ is a strictly increasing function when $x > 0$ and $y > 0$, we immediately have

$$(\xi\eta)_{\sup}(\alpha) = \xi_{\sup}(\alpha)\eta_{\sup}(\alpha), \quad (\xi\eta)_{\inf}(\alpha) = \xi_{\inf}(\alpha)\eta_{\inf}(\alpha) \tag{1.186}$$

by using Theorem 1.52.

Theorem 1.53 *(Zuo [253]) Let $\xi_1, \xi_2, \cdots, \xi_n$ be independent uncertain variables with regular uncertainty distributions. If f is a continuous and strictly decreasing function, then*

$$\xi_{\sup}(\alpha) = f(\xi_{1\inf}(\alpha), \xi_{2\inf}(\alpha), \cdots, \xi_{n\inf}(\alpha)), \qquad (1.187)$$

$$\xi_{\inf}(\alpha) = f(\xi_{1\sup}(\alpha), \xi_{2\sup}(\alpha), \cdots, \xi_{n\sup}(\alpha)). \qquad (1.188)$$

Proof: Since f is a strictly decreasing function, it follows from Theorem 1.30 that the uncertainty distribution Ψ of ξ is determined by

$$\Psi^{-1}(\alpha) = f(\Phi_1^{-1}(1 - \alpha), \Phi_2^{-1}(1 - \alpha), \cdots, \Phi_n^{-1}(1 - \alpha)).$$

By using Theorem 1.46, we get (1.187) and (1.188). The theorem is proved.

Exercise 1.3. What are the critical values to an alternating monotone function of uncertain variables?

Exercise 1.4. Let ξ and η be independent and positive uncertain variables with regular uncertainty distributions. Prove

$$\left(\frac{\xi}{\eta}\right)_{\sup}(\alpha) = \frac{\xi_{\sup}(\alpha)}{\eta_{\inf}(\alpha)}, \qquad \left(\frac{\xi}{\eta}\right)_{\inf}(\alpha) = \frac{\xi_{\inf}(\alpha)}{\eta_{\sup}(\alpha)}. \qquad (1.189)$$

1.11 Entropy

This section provides a definition of entropy to characterize the uncertainty of uncertain variables resulting from information deficiency.

Definition 1.25 *(Liu [130]) Suppose that ξ is an uncertain variable with uncertainty distribution Φ. Then its entropy is defined by*

$$H[\xi] = \int_{-\infty}^{+\infty} S(\Phi(x))\mathrm{d}x \qquad (1.190)$$

where $S(t) = -t\ln t - (1 - t)\ln(1 - t)$.

Example 1.56: Let ξ be an uncertain variable with uncertainty distribution

$$\Phi(x) = \begin{cases} 0, & \text{if } x < a \\ 1, & \text{if } x \geq a. \end{cases} \qquad (1.191)$$

Essentially, ξ is a constant a. It follows from the definition of entropy that

$$H[\xi] = -\int_{-\infty}^{a} (0\ln 0 + 1\ln 1)\,\mathrm{d}x - \int_{a}^{+\infty} (1\ln 1 + 0\ln 0)\,\mathrm{d}x = 0.$$
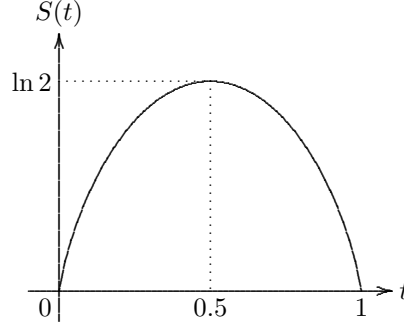
Figure 1.11: Function $S(t) = -t \ln t - (1 - t) \ln(1 - t)$. It is easy to verify that $S(t)$ is a symmetric function about $t = 0.5$, strictly increases on the interval $[0, 0.5]$, strictly decreases on the interval $[0.5, 1]$, and reaches its unique maximum $\ln 2$ at $t = 0.5$.

This means a constant has no uncertainty.

Example 1.57: Let ξ be an uncertain variable on $[a, b]$ with uncertainty distribution

$$\Phi(x) = \begin{cases} 0, & \text{if } x < a \\ 0.5, & \text{if } a \le x < b \\ 1, & \text{if } x \ge b. \end{cases} \tag{1.192}$$

It follows from the definition of entropy that

$$H[\xi] = -\int_a^b (0.5 \ln 0.5 + (1 - 0.5) \ln (1 - 0.5))\, \mathrm{d}x = (b - a) \ln 2.$$

Example 1.58: Let ξ be a linear uncertain variable $\mathcal{L}(a, b)$. Then its entropy is

$$H[\xi] = -\int_a^b \left(\frac{x - a}{b - a} \ln \frac{x - a}{b - a} + \frac{b - x}{b - a} \ln \frac{b - x}{b - a} \right) \mathrm{d}x = \frac{b - a}{2}. \tag{1.193}$$

Example 1.59: Let ξ be a zigzag uncertain variable $\mathcal{Z}(a, b, c)$. Then its entropy is

$$H[\xi] = \frac{c - a}{2}. \tag{1.194}$$

Example 1.60: Let ξ be a normal uncertain variable $\mathcal{N}(e, \sigma)$. Then its entropy is

$$H[\xi] = \frac{\pi \sigma}{\sqrt{3}}. \tag{1.195}$$

Theorem 1.54 *Let ξ be an uncertain variable. Then $H[\xi] \geq 0$ and equality holds if ξ is essentially a constant.*

Proof: The positivity is clear. In addition, when an uncertain variable tends to a constant, its entropy tends to the minimum 0.

Theorem 1.55 *Let ξ be an uncertain variable taking values on the interval $[a, b]$. Then*

$$H[\xi] \leq (b - a) \ln 2 \qquad (1.196)$$

and equality holds if ξ has an uncertainty distribution (1.192).

Proof: The theorem follows from the fact that the function $S(t)$ reaches its maximum $\ln 2$ at $t = 0.5$.

Theorem 1.56 *Let ξ be an uncertain variable, and let c be a real number. Then*

$$H[\xi + c] = H[\xi]. \qquad (1.197)$$

That is, the entropy is invariant under arbitrary translations.

Proof: Write the uncertainty distribution of ξ by Φ. Then the uncertain variable $\xi + c$ has an uncertainty distribution $\Phi(x - c)$. It follows from the definition of entropy that

$$H[\xi + c] = \int_{-\infty}^{+\infty} S\left(\Phi(x - c)\right) \mathrm{d}x = \int_{-\infty}^{+\infty} S(\Phi(x)) \mathrm{d}x = H[\xi].$$

The theorem is proved.

Theorem 1.57 *(Dai and Chen [24]) Let ξ be an uncertain variable with regular uncertainty distribution Φ. Then*

$$H[\xi] = \int_0^1 \Phi^{-1}(\alpha) \ln \frac{\alpha}{1 - \alpha} \mathrm{d}\alpha. \qquad (1.198)$$

Proof: It is clear that $S(\alpha)$ is a derivable function with $S'(\alpha) = -\ln \alpha/(1 - \alpha)$. Since

$$S(\Phi(x)) = \int_0^{\Phi(x)} S'(\alpha) \mathrm{d}\alpha = -\int_{\Phi(x)}^1 S'(\alpha) \mathrm{d}\alpha,$$

we have

$$H[\xi] = \int_{-\infty}^{+\infty} S(\Phi(x)) \mathrm{d}x = \int_{-\infty}^0 \int_0^{\Phi(x)} S'(\alpha) \mathrm{d}\alpha \mathrm{d}x - \int_0^{+\infty} \int_{\Phi(x)}^1 S'(\alpha) \mathrm{d}\alpha \mathrm{d}x.$$

It follows from Fubini theorem that

$$H[\xi] = \int_0^{\Phi(0)} \int_{\Phi^{-1}(\alpha)}^0 S'(\alpha)\mathrm{d}x\mathrm{d}\alpha - \int_{\Phi(0)}^1 \int_0^{\Phi^{-1}(\alpha)} S'(\alpha)\mathrm{d}x\mathrm{d}\alpha$$

$$= -\int_0^{\Phi(0)} \Phi^{-1}(\alpha)S'(\alpha)\mathrm{d}\alpha - \int_{\Phi(0)}^1 \Phi^{-1}(\alpha)S'(\alpha)\mathrm{d}\alpha$$

$$= -\int_0^1 \Phi^{-1}(\alpha)S'(\alpha)\mathrm{d}\alpha = \int_0^1 \Phi^{-1}(\alpha) \ln \frac{\alpha}{1-\alpha}\mathrm{d}\alpha.$$

The theorem is verified.

Theorem 1.58 *(Dai and Chen [24]) Let ξ and η be independent uncertain variables. Then for any real numbers a and b, we have*

$$H[a\xi + b\eta] = |a|H[\xi] + |b|H[\eta]. \tag{1.199}$$

Proof: Without loss of generality, we suppose that ξ and η have regular uncertainty distributions Φ and Ψ, respectively. Otherwise, we may give them a small perturbation such that the uncertainty distributions are regular.

STEP 1: We prove $H[a\xi] = |a|H[\xi]$. If $a > 0$, then the uncertainty distribution Υ of $a\xi$ is determined by

$$\Upsilon^{-1}(\alpha) = a\Phi^{-1}(\alpha).$$

It follows from Theorem 1.57 that

$$H[a\xi] = \int_0^1 a\Phi^{-1}(\alpha) \ln \frac{\alpha}{1-\alpha}\mathrm{d}\alpha = a\int_0^1 \Phi^{-1}(\alpha) \ln \frac{\alpha}{1-\alpha}\mathrm{d}\alpha = |a|H[\xi].$$

If $a = 0$, we immediately have $H[a\xi] = 0 = |a|H[\xi]$. If $a < 0$, then the uncertainty distribution Υ of $a\xi$ is determined by

$$\Upsilon^{-1}(\alpha) = a\Phi^{-1}(1 - \alpha).$$

It follows from Theorem 1.57 that

$$H[a\xi] = \int_0^1 a\Phi^{-1}(1-\alpha) \ln \frac{\alpha}{1-\alpha}\mathrm{d}\alpha = (-a)\int_0^1 \Phi^{-1}(\alpha) \ln \frac{\alpha}{1-\alpha}\mathrm{d}\alpha = |a|H[\xi].$$

Thus we always have $H[a\xi] = |a|H[\xi]$.

STEP 2: We prove $H[\xi + \eta] = H[\xi] + H[\eta]$. Note that the uncertainty distribution Υ of $\xi + \eta$ is determined by

$$\Upsilon^{-1}(\alpha) = \Phi^{-1}(\alpha) + \Psi^{-1}(\alpha).$$

It follows from Theorem 1.57 that

$$H[\xi + \eta] = \int_0^1 (\Phi^{-1}(\alpha) + \Psi^{-1}(\alpha)) \ln \frac{\alpha}{1-\alpha} \mathrm{d}\alpha = H[\xi] + H[\eta].$$

STEP 3: Finally, for any real numbers a and b, it follows from Steps 1 and 2 that

$$H[a\xi + b\eta] = H[a\xi] + H[b\eta] = |a|H[\xi] + |b|H[\eta].$$

The theorem is proved.

Entropy of Function of Uncertain Variables

Theorem 1.59 *(Dai and Chen [24]) Let $\xi_1, \xi_2, \cdots, \xi_n$ be independent uncertain variables with regular uncertainty distributions $\Phi_1, \Phi_2, \cdots, \Phi_n$, respectively. If $f : \Re^n \to \Re$ is a strictly increasing function, then the uncertain variable $\xi = f(\xi_1, \xi_2, \cdots, \xi_n)$ has an entropy*

$$H[\xi] = \int_0^1 f(\Phi_1^{-1}(\alpha), \Phi_2^{-1}(\alpha), \cdots, \Phi_n^{-1}(\alpha)) \ln \frac{\alpha}{1-\alpha} \mathrm{d}\alpha. \qquad (1.200)$$

Proof: Since f is a strictly increasing function, it follows from Theorem 1.25 that the uncertainty distribution Ψ of ξ is determined by

$$\Psi^{-1}(\alpha) = f(\Phi_1^{-1}(\alpha), \Phi_2^{-1}(\alpha), \cdots, \Phi_n^{-1}(\alpha)).$$

By using Theorem 1.57, we get (1.200). The theorem is thus verified.

Exercise 1.5. Let ξ and η be independent and nonnegative uncertain variables with regular uncertainty distributions Φ and Ψ, respectively. Then

$$H[\xi\eta] = \int_0^1 \Phi^{-1}(\alpha)\Psi^{-1}(\alpha) \ln \frac{\alpha}{1-\alpha} \mathrm{d}\alpha. \qquad (1.201)$$

Theorem 1.60 *(Dai and Chen [24]) Let $\xi_1, \xi_2, \cdots, \xi_n$ be independent uncertain variables with regular uncertainty distributions $\Phi_1, \Phi_2, \cdots, \Phi_n$, respectively. If f is a strictly decreasing function, then*

$$H[\xi] = \int_0^1 f(\Phi_1^{-1}(\alpha), \Phi_2^{-1}(\alpha), \cdots, \Phi_n^{-1}(\alpha)) \ln \frac{1-\alpha}{\alpha} \mathrm{d}\alpha. \qquad (1.202)$$

Proof: Since f is a strictly decreasing function, it follows from Theorem 1.30 that the uncertainty distribution Ψ of ξ is determined by

$$\Psi^{-1}(\alpha) = f(\Phi_1^{-1}(1-\alpha), \Phi_2^{-1}(1-\alpha), \cdots, \Phi_n^{-1}(1-\alpha)).$$

By using Theorem 1.57, we get (1.202). The theorem is thus verified.

Exercise 1.6. What is the entropy of an alternating monotone function of uncertain variables?

Exercise 1.7. Let ξ and η be independent and positive uncertain variables with regular uncertainty distributions Φ and Ψ, respectively. Prove

$$H\left[\frac{\xi}{\eta}\right] = \int_0^1 \frac{\Phi^{-1}(\alpha)}{\Psi^{-1}(1-\alpha)} \ln \frac{\alpha}{1-\alpha} \mathrm{d}\alpha. \qquad (1.203)$$

Maximum Entropy Principle

Given some constraints, for example, expected value and variance, there are usually multiple compatible uncertainty distributions. Which uncertainty distribution shall we take? The *maximum entropy principle* attempts to select the uncertainty distribution that maximizes the value of entropy and satisfies the prescribed constraints.

Theorem 1.61 *(Chen and Dai [19]) Let ξ be an uncertain variable whose uncertainty distribution is arbitrary but the expected value e and variance σ^2. Then*

$$H[\xi] \le \frac{\pi\sigma}{\sqrt{3}} \qquad (1.204)$$

and the equality holds if ξ is a normal uncertain variable $\mathcal{N}(e, \sigma)$.

Proof: Let $\Phi(x)$ be the uncertainty distribution of ξ and write $\Psi(x) = \Phi(2e - x)$ for $x \ge e$. It follows from the stipulation (1.168) that the variance is

$$V[\xi] = 2\int_e^{+\infty} (x-e)(1-\Phi(x))\mathrm{d}x + 2\int_e^{+\infty} (x-e)\Psi(x)\mathrm{d}x = \sigma^2.$$

Thus there exists a real number κ such that

$$2\int_e^{+\infty} (x-e)(1-\Phi(x))\mathrm{d}x = \kappa\sigma^2,$$

$$2\int_e^{+\infty} (x-e)\Psi(x)\mathrm{d}x = (1-\kappa)\sigma^2.$$

The maximum entropy distribution Φ should maximize the entropy

$$H[\xi] = \int_{-\infty}^{+\infty} S(\Phi(x))\mathrm{d}x = \int_e^{+\infty} S(\Phi(x))\mathrm{d}x + \int_e^{+\infty} S(\Psi(x))\mathrm{d}x$$

subject to the above two constraints. The Lagrangian is

$$L = \int_e^{+\infty} S(\Phi(x))\mathrm{d}x + \int_e^{+\infty} S(\Psi(x))\mathrm{d}x$$

$$-\alpha\left(2\int_e^{+\infty}(x-e)(1-\Phi(x))\mathrm{d}x - \kappa\sigma^2\right)$$

$$-\beta\left(2\int_e^{+\infty}(x-e)\Psi(x)\mathrm{d}x - (1-\kappa)\sigma^2\right).$$

The maximum entropy distribution meets Euler-Lagrange equations

$$\ln\Phi(x) - \ln(1-\Phi(x)) = 2\alpha(x-e),$$

$$\ln\Psi(x) - \ln(1-\Psi(x)) = 2\beta(e-x).$$

Thus Φ and Ψ have the forms

$$\Phi(x) = (1+\exp(2\alpha(e-x)))^{-1},$$

$$\Psi(x) = (1+\exp(2\beta(x-e)))^{-1}.$$

Substituting them into the variance constraints, we get

$$\Phi(x) = \left(1+\exp\left(\frac{\pi(e-x)}{\sqrt{6}\kappa\sigma}\right)\right)^{-1},$$

$$\Psi(x) = \left(1+\exp\left(\frac{\pi(x-e)}{\sqrt{6(1-\kappa)}\sigma}\right)\right)^{-1}.$$

Then the entropy is

$$H[\xi] = \frac{\pi\sigma\sqrt{\kappa}}{\sqrt{6}} + \frac{\pi\sigma\sqrt{1-\kappa}}{\sqrt{6}}$$

which achieves the maximum when $\kappa = 1/2$. Thus the maximum entropy distribution is just the normal uncertainty distribution $\mathcal{N}(e,\sigma)$.

1.12   Distance

Definition 1.26 *(Liu [127]) The distance between uncertain variables ξ and η is defined as*

$$d(\xi,\eta) = E[|\xi-\eta|]. \tag{1.205}$$

Theorem 1.62 *Let ξ,η,τ be uncertain variables, and let $d(\cdot,\cdot)$ be the distance. Then we have*
(a) (Nonnegativity) $d(\xi,\eta) \geq 0$;
(b) (Identification) $d(\xi,\eta) = 0$ if and only if $\xi = \eta$;
(c) (Symmetry) $d(\xi,\eta) = d(\eta,\xi)$;
(d) (Triangle Inequality) $d(\xi,\eta) \leq 2d(\xi,\tau) + 2d(\eta,\tau)$.

Proof: The parts (a), (b) and (c) follow immediately from the definition. Now we prove the part (d). It follows from the countable subadditivity axiom that

$$
\begin{aligned}
d(\xi, \eta) &= \int_0^{+\infty} \mathcal{M}\left\{|\xi - \eta| \geq r\right\} \mathrm{d}r \\
&\leq \int_0^{+\infty} \mathcal{M}\left\{|\xi - \tau| + |\tau - \eta| \geq r\right\} \mathrm{d}r \\
&\leq \int_0^{+\infty} \mathcal{M}\left\{(|\xi - \tau| \geq r/2) \cup (|\tau - \eta| \geq r/2)\right\} \mathrm{d}r \\
&\leq \int_0^{+\infty} \left(\mathcal{M}\{|\xi - \tau| \geq r/2\} + \mathcal{M}\{|\tau - \eta| \geq r/2\}\right) \mathrm{d}r \\
&= 2E[|\xi - \tau|] + 2E[|\tau - \eta|] = 2d(\xi, \tau) + 2d(\tau, \eta).
\end{aligned}
$$

Example 1.61: Let $\Gamma = \{\gamma_1, \gamma_2, \gamma_3\}$. Define $\mathcal{M}\{\emptyset\} = 0$, $\mathcal{M}\{\Gamma\} = 1$ and $\mathcal{M}\{\Lambda\} = 1/2$ for any subset Λ (excluding $\emptyset$ and Γ). We set uncertain variables ξ, η and τ as follows,

$$
\xi(\gamma) = \left\{ \begin{array}{ll} 1, & \text{if } \gamma = \gamma_1 \\ 1, & \text{if } \gamma = \gamma_2 \\ 0, & \text{if } \gamma = \gamma_3, \end{array} \right. \quad \eta(\gamma) = \left\{ \begin{array}{ll} 0, & \text{if } \gamma = \gamma_1 \\ -1, & \text{if } \gamma = \gamma_2 \\ -1, & \text{if } \gamma = \gamma_3, \end{array} \right. \quad \tau(\gamma) \equiv 0.
$$

It is easy to verify that $d(\xi, \tau) = d(\tau, \eta) = 1/2$ and $d(\xi, \eta) = 3/2$. Thus

$$
d(\xi, \eta) = \frac{3}{2}(d(\xi, \tau) + d(\tau, \eta)).
$$

A conjecture is $d(\xi, \eta) \leq 1.5(d(\xi, \tau) + d(\tau, \eta))$ for arbitrary uncertain variables ξ, η and τ. This is an open problem.

1.13 Inequalities

Theorem 1.63 *(Liu [127]) Let ξ be an uncertain variable, and f a non-negative function. If f is even and increasing on $[0, \infty)$, then for any given number $t > 0$, we have*

$$
\mathcal{M}\{|\xi| \geq t\} \leq \frac{E[f(\xi)]}{f(t)}. \tag{1.206}
$$

Proof: It is clear that $\mathcal{M}\{|\xi| \geq f^{-1}(r)\}$ is a monotone decreasing function of r on $[0, \infty)$. It follows from the nonnegativity of $f(\xi)$ that

$$
\begin{aligned}
E[f(\xi)] &= \int_0^{+\infty} \mathcal{M}\{f(\xi) \geq r\}\mathrm{d}r = \int_0^{+\infty} \mathcal{M}\{|\xi| \geq f^{-1}(r)\}\mathrm{d}r \\
&\geq \int_0^{f(t)} \mathcal{M}\{|\xi| \geq f^{-1}(r)\}\mathrm{d}r \geq \int_0^{f(t)} \mathrm{d}r \cdot \mathcal{M}\{|\xi| \geq f^{-1}(f(t))\} \\
&= f(t) \cdot \mathcal{M}\{|\xi| \geq t\}
\end{aligned}
$$

which proves the inequality.

Theorem 1.64 *(Liu [127], Markov Inequality) Let ξ be an uncertain variable. Then for any given numbers $t > 0$ and $p > 0$, we have*

$$
\mathcal{M}\{|\xi| \geq t\} \leq \frac{E[|\xi|^p]}{t^p}. \tag{1.207}
$$

Proof: It is a special case of Theorem 1.63 when $f(x) = |x|^p$.

Example 1.62: For any given positive number t, we define an uncertain variable as follows,

$$
\xi = \begin{cases} 0 \text{ with uncertain measure } 1/2 \\ t \text{ with uncertain measure } 1/2. \end{cases}
$$

Then $E[\xi^p] = t^p/2$ and $\mathcal{M}\{\xi \geq t\} = 1/2 = E[\xi^p]/t^p$.

Theorem 1.65 *(Liu [127], Chebyshev Inequality) Let ξ be an uncertain variable whose variance $V[\xi]$ exists. Then for any given number $t > 0$, we have*

$$
\mathcal{M}\{|\xi - E[\xi]| \geq t\} \leq \frac{V[\xi]}{t^2}. \tag{1.208}
$$

Proof: It is a special case of Theorem 1.63 when the uncertain variable ξ is replaced with $\xi - E[\xi]$, and $f(x) = x^2$.

Example 1.63: For any given positive number t, we define an uncertain variable as follows,

$$
\xi = \begin{cases} -t \text{ with uncertain measure } 1/2 \\ t \text{ with uncertain measure } 1/2. \end{cases}
$$

Then $V[\xi] = t^2$ and $\mathcal{M}\{|\xi - E[\xi]| \geq t\} = 1 = V[\xi]/t^2$.

Theorem 1.66 *(Liu [127], Hölder's Inequality) Let p and q be positive numbers with $1/p + 1/q = 1$, and let ξ and η be independent uncertain variables with $E[|\xi|^p] < \infty$ and $E[|\eta|^q] < \infty$. Then we have*

$$
E[|\xi\eta|] \leq \sqrt[p]{E[|\xi|^p]}\sqrt[q]{E[|\eta|^q]}. \tag{1.209}
$$

Proof: The inequality holds trivially if at least one of ξ and η is zero a.s. Now we assume $E[|\xi|^p] > 0$ and $E[|\eta|^q] > 0$. It is easy to prove that the function $f(x, y) = \sqrt[p]{x}\sqrt[q]{y}$ is a concave function on $\{(x, y) : x \geq 0, y \geq 0\}$. Thus for any point (x_0, y_0) with $x_0 > 0$ and $y_0 > 0$, there exist two real numbers a and b such that

$$f(x, y) - f(x_0, y_0) \leq a(x - x_0) + b(y - y_0), \quad \forall x \geq 0, y \geq 0.$$

Letting $x_0 = E[|\xi|^p]$, $y_0 = E[|\eta|^q]$, $x = |\xi|^p$ and $y = |\eta|^q$, we have

$$f(|\xi|^p, |\eta|^q) - f(E[|\xi|^p], E[|\eta|^q]) \leq a(|\xi|^p - E[|\xi|^p]) + b(|\eta|^q - E[|\eta|^q]).$$

Taking the expected values on both sides, we obtain

$$E[f(|\xi|^p, |\eta|^q)] \leq f(E[|\xi|^p], E[|\eta|^q]).$$

Hence the inequality (1.209) holds.

Theorem 1.67 *(Liu [127], Minkowski Inequality) Let p be a real number with $p \geq 1$, and let ξ and η be independent uncertain variables with $E[|\xi|^p] < \infty$ and $E[|\eta|^p] < \infty$. Then we have*

$$\sqrt[p]{E[|\xi + \eta|^p]} \leq \sqrt[p]{E[|\xi|^p]} + \sqrt[p]{E[|\eta|^p]}. \tag{1.210}$$

Proof: The inequality holds trivially if at least one of ξ and η is zero a.s. Now we assume $E[|\xi|^p] > 0$ and $E[|\eta|^p] > 0$. It is easy to prove that the function $f(x, y) = (\sqrt[p]{x} + \sqrt[p]{y})^p$ is a concave function on $\{(x, y) : x \geq 0, y \geq 0\}$. Thus for any point (x_0, y_0) with $x_0 > 0$ and $y_0 > 0$, there exist two real numbers a and b such that

$$f(x, y) - f(x_0, y_0) \leq a(x - x_0) + b(y - y_0), \quad \forall x \geq 0, y \geq 0.$$

Letting $x_0 = E[|\xi|^p]$, $y_0 = E[|\eta|^p]$, $x = |\xi|^p$ and $y = |\eta|^p$, we have

$$f(|\xi|^p, |\eta|^p) - f(E[|\xi|^p], E[|\eta|^p]) \leq a(|\xi|^p - E[|\xi|^p]) + b(|\eta|^p - E[|\eta|^p]).$$

Taking the expected values on both sides, we obtain

$$E[f(|\xi|^p, |\eta|^p)] \leq f(E[|\xi|^p], E[|\eta|^p]).$$

Hence the inequality (1.210) holds.

Theorem 1.68 *(Liu [127], Jensen's Inequality) Let ξ be an uncertain variable, and $f : \Re \to \Re$ a convex function. If $E[\xi]$ and $E[f(\xi)]$ are finite, then*

$$f(E[\xi]) \leq E[f(\xi)]. \tag{1.211}$$

Especially, when $f(x) = |x|^p$ and $p \geq 1$, we have $|E[\xi]|^p \leq E[|\xi|^p]$.

Proof: Since f is a convex function, for each y, there exists a number k such that $f(x) - f(y) \geq k \cdot (x - y)$. Replacing x with ξ and y with $E[\xi]$, we obtain

$$f(\xi) - f(E[\xi]) \geq k \cdot (\xi - E[\xi]).$$

Taking the expected values on both sides, we have

$$E[f(\xi)] - f(E[\xi]) \geq k \cdot (E[\xi] - E[\xi]) = 0$$

which proves the inequality.

1.14   Convergence Concepts

We have the following four convergence concepts of uncertain sequence: convergence almost surely (a.s.), convergence in measure, convergence in mean, and convergence in distribution.

Table 1.1: Relationship among Convergence Concepts

Convergence in Mean	$\Rightarrow$	Convergence in Measure	$\Rightarrow$	Convergence in Distribution

Definition 1.27 *(Liu [127]) Suppose that $\xi, \xi_1, \xi_2, \cdots$ are uncertain variables defined on the uncertainty space $(\Gamma, \mathcal{L}, \mathcal{M})$. The sequence $\{\xi_i\}$ is said to be convergent a.s. to ξ if there exists an event Λ with $\mathcal{M}\{\Lambda\} = 1$ such that*

$$\lim_{i \to \infty} |\xi_i(\gamma) - \xi(\gamma)| = 0 \tag{1.212}$$

for every $\gamma \in \Lambda$. In that case we write $\xi_i \to \xi$, a.s.

Definition 1.28 *(Liu [127]) Suppose that $\xi, \xi_1, \xi_2, \cdots$ are uncertain variables. We say that the sequence $\{\xi_i\}$ converges in measure to ξ if*

$$\lim_{i \to \infty} \mathcal{M}\{|\xi_i - \xi| \geq \varepsilon\} = 0 \tag{1.213}$$

for every $\varepsilon > 0$.

Definition 1.29 *(Liu [127]) Suppose that $\xi, \xi_1, \xi_2, \cdots$ are uncertain variables with finite expected values. We say that the sequence $\{\xi_i\}$ converges in mean to ξ if*

$$\lim_{i \to \infty} E[|\xi_i - \xi|] = 0. \tag{1.214}$$

In addition, the sequence $\{\xi_i\}$ is said to converge in mean square to ξ if

$$\lim_{i \to \infty} E[|\xi_i - \xi|^2] = 0. \tag{1.215}$$

Definition 1.30 *(Liu [127]) Suppose that $\Phi, \Phi_1, \Phi_2, \cdots$ are the uncertainty distributions of uncertain variables $\xi, \xi_1, \xi_2, \cdots$, respectively. We say that $\{\xi_i\}$ converges in distribution to ξ if $\Phi_i \to \Phi$ at any continuity point of Φ.*

Convergence in Mean vs. Convergence in Measure

Theorem 1.31 *(Liu [127]) Suppose that $\xi, \xi_1, \xi_2, \cdots$ are uncertain variables. If $\{\xi_i\}$ converges in mean to ξ, then $\{\xi_i\}$ converges in measure to ξ.*

Proof: It follows from the Markov inequality that for any given number $\varepsilon > 0$, we have

$$\mathcal{M}\{|\xi_i - \xi| \geq \varepsilon\} \leq \frac{E[|\xi_i - \xi|]}{\varepsilon} \to 0$$

as $i \to \infty$. Thus $\{\xi_i\}$ converges in measure to ξ. The theorem is proved.

Example 1.64: Convergence in measure does not imply convergence in mean. Take an uncertainty space $(\Gamma, \mathcal{L}, \mathcal{M})$ to be $\{\gamma_1, \gamma_2, \cdots\}$ with

$$\mathcal{M}\{\Lambda\} = \begin{cases} \sup\limits_{\gamma_i \in \Lambda} 1/i, & \text{if } \sup\limits_{\gamma_i \in \Lambda} 1/i < 0.5 \\ 1 - \sup\limits_{\gamma_i \notin \Lambda} 1/i, & \text{if } \sup\limits_{\gamma_i \notin \Lambda} 1/i < 0.5 \\ 0.5, & \text{otherwise.} \end{cases}$$

The uncertain variables are defined by

$$\xi_i(\gamma_j) = \begin{cases} i, & \text{if } j = i \\ 0, & \text{otherwise} \end{cases}$$

for $i = 1, 2, \cdots$ and $\xi \equiv 0$. For some small number $\varepsilon > 0$, we have

$$\mathcal{M}\{|\xi_i - \xi| \geq \varepsilon\} = \mathcal{M}\{|\xi_i - \xi| \geq \varepsilon\} = \frac{1}{i} \to 0.$$

That is, the sequence $\{\xi_i\}$ converges in measure to ξ. However, for each i, we have

$$E[|\xi_i - \xi|] = 1.$$

That is, the sequence $\{\xi_i\}$ does not converge in mean to ξ.

Convergence in Measure vs. Convergence in Distribution

Theorem 1.32 *(Liu [127]) Suppose $\xi, \xi_1, \xi_2, \cdots$ are uncertain variables. If $\{\xi_i\}$ converges in measure to ξ, then $\{\xi_i\}$ converges in distribution to ξ.*

Proof: Let x be a given continuity point of the uncertainty distribution Φ. On the one hand, for any $y > x$, we have

$$\{\xi_i \leq x\} = \{\xi_i \leq x, \xi \leq y\} \cup \{\xi_i \leq x, \xi > y\} \subset \{\xi \leq y\} \cup \{|\xi_i - \xi| \geq y - x\}.$$

It follows from the countable subadditivity axiom that

$$\Phi_i(x) \leq \Phi(y) + \mathcal{M}\{|\xi_i - \xi| \geq y - x\}.$$

Since $\{\xi_i\}$ converges in measure to ξ, we have $\mathcal{M}\{|\xi_i - \xi| \geq y - x\} \to 0$ as $i \to \infty$. Thus we obtain $\limsup_{i\to\infty} \Phi_i(x) \leq \Phi(y)$ for any $y > x$. Letting $y \to x$, we get

$$\limsup_{i\to\infty} \Phi_i(x) \leq \Phi(x). \tag{1.216}$$

On the other hand, for any $z < x$, we have

$$\{\xi \leq z\} = \{\xi_i \leq x, \xi \leq z\} \cup \{\xi_i > x, \xi \leq z\} \subset \{\xi_i \leq x\} \cup \{|\xi_i - \xi| \geq x - z\}$$

which implies that

$$\Phi(z) \leq \Phi_i(x) + \mathcal{M}\{|\xi_i - \xi| \geq x - z\}.$$

Since $\mathcal{M}\{|\xi_i - \xi| \geq x - z\} \to 0$, we obtain $\Phi(z) \leq \liminf_{i\to\infty} \Phi_i(x)$ for any $z < x$. Letting $z \to x$, we get

$$\Phi(x) \leq \liminf_{i\to\infty} \Phi_i(x). \tag{1.217}$$

It follows from (1.216) and (1.217) that $\Phi_i(x) \to \Phi(x)$. The theorem is proved.

Example 1.65: Convergence in distribution does not imply convergence in measure. Take an uncertainty space $(\Gamma, \mathcal{L}, \mathcal{M})$ to be $\{\gamma_1, \gamma_2\}$ with $\mathcal{M}\{\gamma_1\} = \mathcal{M}\{\gamma_2\} = 1/2$. We define an uncertain variables as

$$\xi(\gamma) = \begin{cases} -1, & \text{if } \gamma = \gamma_1 \\ 1, & \text{if } \gamma = \gamma_2. \end{cases}$$

We also define $\xi_i = -\xi$ for $i = 1, 2, \cdots$ Then ξ_i and ξ have the same chance distribution. Thus $\{\xi_i\}$ converges in distribution to ξ. However, for some small number $\varepsilon > 0$, we have

$$\mathcal{M}\{|\xi_i - \xi| \geq \varepsilon\} = \mathcal{M}\{|\xi_i - \xi| \geq \varepsilon\} = 1.$$

That is, the sequence $\{\xi_i\}$ does not converge in measure to ξ.

Convergence Almost Surely vs. Convergence in Measure

Example 1.66: Convergence a.s. does not imply convergence in measure. Take an uncertainty space $(\Gamma, \mathcal{L}, \mathcal{M})$ to be $\{\gamma_1, \gamma_2, \cdots\}$ with

$$\mathcal{M}\{\Lambda\} = \begin{cases} \sup_{\gamma_i \in \Lambda} i/(2i+1), & \text{if } \sup_{\gamma_i \in \Lambda} i/(2i+1) < 0.5 \\ 1 - \sup_{\gamma_i \notin \Lambda} i/(2i+1), & \text{if } \sup_{\gamma_i \notin \Lambda} i/(2i+1) < 0.5 \\ 0.5, & \text{otherwise.} \end{cases}$$

Then we define uncertain variables as

$$\xi_i(\gamma_j) = \begin{cases} i, & \text{if } j = i \\ 0, & \text{otherwise} \end{cases}$$

for $i = 1, 2, \cdots$ and $\xi \equiv 0$. The sequence $\{\xi_i\}$ converges a.s. to ξ. However, for some small number $\varepsilon > 0$, we have

$$\mathcal{M}\{|\xi_i - \xi| \geq \varepsilon\} = \mathcal{M}\{|\xi_i - \xi| \geq \varepsilon\} = \frac{i}{2i+1} \to \frac{1}{2}.$$

That is, the sequence $\{\xi_i\}$ does not converge in measure to ξ.

Example 1.67: Convergence in measure does not imply convergence a.s. Take an uncertainty space $(\Gamma, \mathcal{L}, \mathcal{M})$ to be $[0, 1]$ with Borel algebra and Lebesgue measure. For any positive integer i, there is an integer j such that $i = 2^j + k$, where k is an integer between 0 and $2^j - 1$. Then we define uncertain variables as

$$\xi_i(\gamma) = \begin{cases} 1, & \text{if } k/2^j \leq \gamma \leq (k+1)/2^j \\ 0, & \text{otherwise} \end{cases}$$

for $i = 1, 2, \cdots$ and $\xi \equiv 0$. For some small number $\varepsilon > 0$, we have

$$\mathcal{M}\{|\xi_i - \xi| \geq \varepsilon\} = \mathcal{M}\{|\xi_i - \xi| \geq \varepsilon\} = \frac{1}{2^j} \to 0$$

as $i \to \infty$. That is, the sequence $\{\xi_i\}$ converges in measure to ξ. However, for any $\gamma \in [0, 1]$, there is an infinite number of intervals of the form $[k/2^j, (k+1)/2^j]$ containing γ. Thus $\xi_i(\gamma)$ does not converge to 0. In other words, the sequence $\{\xi_i\}$ does not converge a.s. to ξ.

Convergence Almost Surely vs. Convergence in Mean

Example 1.68: Convergence a.s. does not imply convergence in mean. Take an uncertainty space $(\Gamma, \mathcal{L}, \mathcal{M})$ to be $\{\gamma_1, \gamma_2, \cdots\}$ with

$$\mathcal{M}\{\Lambda\} = \sum_{\gamma_i \in \Lambda} \frac{1}{2^i}.$$

The uncertain variables are defined by

$$\xi_i(\gamma_j) = \begin{cases} 2^i, & \text{if } j = i \\ 0, & \text{otherwise} \end{cases}$$

for $i = 1, 2, \cdots$ and $\xi \equiv 0$. Then ξ_i converges a.s. to ξ. However, the sequence $\{\xi_i\}$ does not converge in mean to ξ because $E[|\xi_i - \xi|] \equiv 1$.

Example 1.69: Convergence in mean does not imply convergence a.s. Take an uncertainty space $(\Gamma, \mathcal{L}, \mathcal{M})$ to be $[0, 1]$ with Borel algebra and Lebesgue measure. For any positive integer i, there is an integer j such that $i = 2^j + k$, where k is an integer between 0 and $2^j - 1$. The uncertain variables are defined by

$$\xi_i(\gamma) = \begin{cases} 1, & \text{if } k/2^j \leq \gamma \leq (k+1)/2^j \\ 0, & \text{otherwise} \end{cases}$$

for $i = 1, 2, \cdots$ and $\xi \equiv 0$. Then

$$E[|\xi_i - \xi|] = \frac{1}{2^j} \to 0.$$

That is, the sequence $\{\xi_i\}$ converges in mean to ξ. However, for any $\gamma \in [0, 1]$, there is an infinite number of intervals of the form $[k/2^j, (k+1)/2^j]$ containing γ. Thus $\xi_i(\gamma)$ does not converge to 0. In other words, the sequence $\{\xi_i\}$ does not converge a.s. to ξ.

Convergence Almost Surely vs. Convergence in Distribution

Example 1.70: Convergence in distribution does not imply convergence a.s. Take an uncertainty space to be $(\Gamma, \mathcal{L}, \mathcal{M})$ to be $\{\gamma_1, \gamma_2\}$ with $\mathcal{M}\{\gamma_1\} = \mathcal{M}\{\gamma_2\} = 1/2$. We define an uncertain variable ξ as

$$\xi(\gamma) = \begin{cases} -1, & \text{if } \gamma = \gamma_1 \\ 1, & \text{if } \gamma = \gamma_2. \end{cases}$$

We also define $\xi_i = -\xi$ for $i = 1, 2, \cdots$ Then ξ_i and ξ have the same uncertainty distribution. Thus $\{\xi_i\}$ converges in distribution to ξ. However, the sequence $\{\xi_i\}$ does not converge a.s. to ξ.

Example 1.71: Convergence a.s. does not imply convergence in distribution. Take an uncertainty space $(\Gamma, \mathcal{L}, \mathcal{M})$ to be $\{\gamma_1, \gamma_2, \cdots\}$ with

$$\mathcal{M}\{\Lambda\} = \begin{cases} \sup\limits_{\gamma_i \in \Lambda} i/(2i+1), & \text{if } \sup\limits_{\gamma_i \in \Lambda} i/(2i+1) < 0.5 \\ 1 - \sup\limits_{\gamma_i \notin \Lambda} i/(2i+1), & \text{if } \sup\limits_{\gamma_i \notin \Lambda} i/(2i+1) < 0.5 \\ 0.5, & \text{otherwise.} \end{cases}$$

The uncertain variables are defined by

$$\xi_i(\gamma_j) = \begin{cases} i, & \text{if } j = i \\ 0, & \text{otherwise} \end{cases}$$

for $i = 1, 2, \cdots$ and $\xi \equiv 0$. Then the sequence $\{\xi_i\}$ converges a.s. to ξ. However, the uncertainty distributions of ξ_i are

$$\Phi_i(x) = \begin{cases} 0, & \text{if } x < 0 \\ (i+1)/(2i+1), & \text{if } 0 \le x < i \\ 1, & \text{if } x \ge i \end{cases}$$

for $i = 1, 2, \cdots$, respectively. The uncertainty distribution of ξ is

$$\Phi(x) = \begin{cases} 0, & \text{if } x < 0 \\ 1, & \text{if } x \ge 0. \end{cases}$$

It is clear that $\Phi_i(x)$ does not converge to $\Phi(x)$ at $x > 0$. That is, the sequence $\{\xi_i\}$ does not converge in distribution to ξ.

1.15   Conditional Uncertainty

We consider the uncertain measure of an event A after it has been learned that some other event B has occurred. This new uncertain measure of A is called the conditional uncertain measure of A given B.

In order to define a conditional uncertain measure $\mathcal{M}\{A|B\}$, at first we have to enlarge $\mathcal{M}\{A \cap B\}$ because $\mathcal{M}\{A \cap B\} < 1$ for all events whenever $\mathcal{M}\{B\} < 1$. It seems that we have no alternative but to divide $\mathcal{M}\{A \cap B\}$ by $\mathcal{M}\{B\}$. Unfortunately, $\mathcal{M}\{A \cap B\}/\mathcal{M}\{B\}$ is not always an uncertain measure. However, the value $\mathcal{M}\{A|B\}$ should not be greater than $\mathcal{M}\{A \cap B\}/\mathcal{M}\{B\}$ (otherwise the normality will be lost), i.e.,

$$\mathcal{M}\{A|B\} \le \frac{\mathcal{M}\{A \cap B\}}{\mathcal{M}\{B\}}. \tag{1.218}$$

On the other hand, in order to preserve the self-duality, we should have

$$\mathcal{M}\{A|B\} = 1 - \mathcal{M}\{A^c|B\} \ge 1 - \frac{\mathcal{M}\{A^c \cap B\}}{\mathcal{M}\{B\}}. \tag{1.219}$$

Furthermore, since $(A \cap B) \cup (A^c \cap B) = B$, we have $\mathcal{M}\{B\} \le \mathcal{M}\{A \cap B\} + \mathcal{M}\{A^c \cap B\}$ by using the countable subadditivity axiom. Thus

$$0 \le 1 - \frac{\mathcal{M}\{A^c \cap B\}}{\mathcal{M}\{B\}} \le \frac{\mathcal{M}\{A \cap B\}}{\mathcal{M}\{B\}} \le 1. \tag{1.220}$$

Hence any numbers between $1 - \mathcal{M}\{A^c \cap B\}/\mathcal{M}\{B\}$ and $\mathcal{M}\{A \cap B\}/\mathcal{M}\{B\}$ are reasonable values that the conditional uncertain measure may take. Based on the maximum uncertainty principle, we have the following conditional uncertain measure.

Definition 1.33 *(Liu [127]) Let* $(\Gamma, \mathcal{L}, \mathcal{M})$ *be an uncertainty space, and* $A, B \in \mathcal{L}$*. Then the conditional uncertain measure of* A *given* B *is defined by*

$$
\mathcal{M}\{A|B\} = \begin{cases} \dfrac{\mathcal{M}\{A \cap B\}}{\mathcal{M}\{B\}}, & \text{if } \dfrac{\mathcal{M}\{A \cap B\}}{\mathcal{M}\{B\}} < 0.5 \\[3mm] 1 - \dfrac{\mathcal{M}\{A^c \cap B\}}{\mathcal{M}\{B\}}, & \text{if } \dfrac{\mathcal{M}\{A^c \cap B\}}{\mathcal{M}\{B\}} < 0.5 \\[3mm] 0.5, & \text{otherwise} \end{cases} \qquad (1.221)
$$

provided that $\mathcal{M}\{B\} > 0$.

It follows immediately from the definition of conditional uncertain measure that

$$
1 - \frac{\mathcal{M}\{A^c \cap B\}}{\mathcal{M}\{B\}} \leq \mathcal{M}\{A|B\} \leq \frac{\mathcal{M}\{A \cap B\}}{\mathcal{M}\{B\}}. \qquad (1.222)
$$

Furthermore, the conditional uncertain measure obeys the maximum uncertainty principle, and takes values as close to 0.5 as possible.

Remark 1.7: Assume that we know the *prior* uncertain measures $\mathcal{M}\{B\}$, $\mathcal{M}\{A \cap B\}$ and $\mathcal{M}\{A^c \cap B\}$. Then the conditional uncertain measure $\mathcal{M}\{A|B\}$ yields the *posterior* uncertain measure of A after the occurrence of event B.

Theorem 1.69 *Let* $(\Gamma, \mathcal{L}, \mathcal{M})$ *be an uncertainty space, and* B *an event with* $\mathcal{M}\{B\} > 0$. *Then* $\mathcal{M}\{\cdot|B\}$ *defined by (1.221) is an uncertain measure, and* $(\Gamma, \mathcal{L}, \mathcal{M}\{\cdot|B\})$ *is an uncertainty space.*

Proof: It is sufficient to prove that $\mathcal{M}\{\cdot|B\}$ satisfies the normality, monotonicity, self-duality and countable subadditivity axioms. At first, it satisfies the normality axiom, i.e.,

$$
\mathcal{M}\{\Gamma|B\} = 1 - \frac{\mathcal{M}\{\Gamma^c \cap B\}}{\mathcal{M}\{B\}} = 1 - \frac{\mathcal{M}\{\emptyset\}}{\mathcal{M}\{B\}} = 1.
$$

For any events A_1 and A_2 with $A_1 \subset A_2$, if

$$
\frac{\mathcal{M}\{A_1 \cap B\}}{\mathcal{M}\{B\}} \leq \frac{\mathcal{M}\{A_2 \cap B\}}{\mathcal{M}\{B\}} < 0.5,
$$

then

$$
\mathcal{M}\{A_1|B\} = \frac{\mathcal{M}\{A_1 \cap B\}}{\mathcal{M}\{B\}} \leq \frac{\mathcal{M}\{A_2 \cap B\}}{\mathcal{M}\{B\}} = \mathcal{M}\{A_2|B\}.
$$

If

$$
\frac{\mathcal{M}\{A_1 \cap B\}}{\mathcal{M}\{B\}} \leq 0.5 \leq \frac{\mathcal{M}\{A_2 \cap B\}}{\mathcal{M}\{B\}},
$$

then $\mathcal{M}\{A_1|B\} \leq 0.5 \leq \mathcal{M}\{A_2|B\}$. If

$$
0.5 < \frac{\mathcal{M}\{A_1 \cap B\}}{\mathcal{M}\{B\}} \leq \frac{\mathcal{M}\{A_2 \cap B\}}{\mathcal{M}\{B\}},
$$

then we have

$$\mathcal{M}\{A_1|B\} = \left(1 - \frac{\mathcal{M}\{A_1^c \cap B\}}{\mathcal{M}\{B\}}\right) \vee 0.5 \leq \left(1 - \frac{\mathcal{M}\{A_2^c \cap B\}}{\mathcal{M}\{B\}}\right) \vee 0.5 = \mathcal{M}\{A_2|B\}.$$

This means that $\mathcal{M}\{\cdot|B\}$ satisfies the monotonicity axiom. For any event A, if

$$\frac{\mathcal{M}\{A \cap B\}}{\mathcal{M}\{B\}} \geq 0.5, \quad \frac{\mathcal{M}\{A^c \cap B\}}{\mathcal{M}\{B\}} \geq 0.5,$$

then we have $\mathcal{M}\{A|B\} + \mathcal{M}\{A^c|B\} = 0.5 + 0.5 = 1$ immediately. Otherwise, without loss of generality, suppose

$$\frac{\mathcal{M}\{A \cap B\}}{\mathcal{M}\{B\}} < 0.5 < \frac{\mathcal{M}\{A^c \cap B\}}{\mathcal{M}\{B\}},$$

then we have

$$\mathcal{M}\{A|B\} + \mathcal{M}\{A^c|B\} = \frac{\mathcal{M}\{A \cap B\}}{\mathcal{M}\{B\}} + \left(1 - \frac{\mathcal{M}\{A \cap B\}}{\mathcal{M}\{B\}}\right) = 1.$$

That is, $\mathcal{M}\{\cdot|B\}$ satisfies the self-duality axiom. Finally, for any countable sequence $\{A_i\}$ of events, if $\mathcal{M}\{A_i|B\} < 0.5$ for all i, it follows from the countable subadditivity axiom that

$$\mathcal{M}\left\{\bigcup_{i=1}^{\infty} A_i \cap B\right\} \leq \frac{\mathcal{M}\left\{\bigcup_{i=1}^{\infty} A_i \cap B\right\}}{\mathcal{M}\{B\}} \leq \frac{\displaystyle\sum_{i=1}^{\infty} \mathcal{M}\{A_i \cap B\}}{\mathcal{M}\{B\}} = \sum_{i=1}^{\infty} \mathcal{M}\{A_i|B\}.$$

Suppose there is one term greater than 0.5, say

$$\mathcal{M}\{A_1|B\} \geq 0.5, \quad \mathcal{M}\{A_i|B\} < 0.5, \quad i = 2, 3, \cdots$$

If $\mathcal{M}\{\cup_i A_i|B\} = 0.5$, then we immediately have

$$\mathcal{M}\left\{\bigcup_{i=1}^{\infty} A_i \cap B\right\} \leq \sum_{i=1}^{\infty} \mathcal{M}\{A_i|B\}.$$

If $\mathcal{M}\{\cup_i A_i|B\} > 0.5$, we may prove the above inequality by the following facts:

$$A_1^c \cap B \subset \bigcup_{i=2}^{\infty}(A_i \cap B) \cup \left(\bigcap_{i=1}^{\infty} A_i^c \cap B\right),$$

$$\mathcal{M}\{A_1^c \cap B\} \leq \sum_{i=2}^{\infty} \mathcal{M}\{A_i \cap B\} + \mathcal{M}\left\{\bigcap_{i=1}^{\infty} A_i^c \cap B\right\},$$

$$\mathcal{M}\left\{\bigcup_{i=1}^{\infty} A_i | B\right\} = 1 - \frac{\mathcal{M}\left\{\bigcap_{i=1}^{\infty} A_i^c \cap B\right\}}{\mathcal{M}\{B\}},$$

$$\sum_{i=1}^{\infty} \mathcal{M}\{A_i | B\} \geq 1 - \frac{\mathcal{M}\{A_1^c \cap B\}}{\mathcal{M}\{B\}} + \frac{\sum_{i=2}^{\infty} \mathcal{M}\{A_i \cap B\}}{\mathcal{M}\{B\}}.$$

If there are at least two terms greater than 0.5, then the countable subadditivity is clearly true. Thus $\mathcal{M}\{\cdot|B\}$ satisfies the countable subadditivity axiom. Hence $\mathcal{M}\{\cdot|B\}$ is an uncertain measure. Furthermore, $(\Gamma, \mathcal{L}, \mathcal{M}\{\cdot|B\})$ is an uncertainty space.

Definition 1.34 *(Liu [127]) Let ξ be an uncertain variable on $(\Gamma, \mathcal{L}, \mathcal{M})$. A conditional uncertain variable of ξ given B is a measurable function $\xi|_B$ from the conditional uncertainty space $(\Gamma, \mathcal{L}, \mathcal{M}\{\cdot|B\})$ to the set of real numbers such that*

$$\xi|_B(\gamma) \equiv \xi(\gamma), \quad \forall \gamma \in \Gamma. \tag{1.223}$$

Definition 1.35 *(Liu [127]) The conditional uncertainty distribution $\Phi\colon \Re \to [0,1]$ of an uncertain variable ξ given B is defined by*

$$\Phi(x|B) = \mathcal{M}\{\xi \leq x|B\} \tag{1.224}$$

provided that $\mathcal{M}\{B\} > 0$.

Theorem 1.70 *Let ξ be an uncertain variable with regular uncertainty distribution $\Phi(x)$, and t a real number with $\Phi(t) < 1$. Then the conditional uncertainty distribution of ξ given $\xi > t$ is*

$$\Phi(x|(t, +\infty)) = \begin{cases} 0, & \text{if } \Phi(x) \leq \Phi(t) \\[2mm] \dfrac{\Phi(x)}{1 - \Phi(t)} \wedge 0.5, & \text{if } \Phi(t) < \Phi(x) \leq (1 + \Phi(t))/2 \\[2mm] \dfrac{\Phi(x) - \Phi(t)}{1 - \Phi(t)}, & \text{if } (1 + \Phi(t))/2 \leq \Phi(x). \end{cases}$$

Proof: It follows from $\Phi(x|(t, +\infty)) = \mathcal{M}\{\xi \leq x|\xi > t\}$ and the definition of conditional uncertainty that

$$\Phi(x|(t, +\infty)) = \begin{cases} \dfrac{\mathcal{M}\{(\xi \leq x) \cap (\xi > t)\}}{\mathcal{M}\{\xi > t\}}, & \text{if } \dfrac{\mathcal{M}\{(\xi \leq x) \cap (\xi > t)\}}{\mathcal{M}\{\xi > t\}} < 0.5 \\[3mm] 1 - \dfrac{\mathcal{M}\{(\xi > x) \cap (\xi > t)\}}{\mathcal{M}\{\xi > t\}}, & \text{if } \dfrac{\mathcal{M}\{(\xi > x) \cap (\xi > t)\}}{\mathcal{M}\{\xi > t\}} < 0.5 \\[3mm] 0.5, & \text{otherwise.} \end{cases}$$

When $\Phi(x) \leq \Phi(t)$, we have $x \leq t$, and

$$\frac{\mathcal{M}\{(\xi \leq x) \cap (\xi > t)\}}{\mathcal{M}\{\xi > t\}} = \frac{\mathcal{M}\{\emptyset\}}{1 - \Phi(t)} = 0 < 0.5.$$

Thus

$$\Phi(x|(t, +\infty)) = \frac{\mathcal{M}\{(\xi \leq x) \cap (\xi > t)\}}{\mathcal{M}\{\xi > t\}} = 0.$$

When $\Phi(t) < \Phi(x) \leq (1 + \Phi(t))/2$, we have $x > t$, and

$$\frac{\mathcal{M}\{(\xi > x) \cap (\xi > t)\}}{\mathcal{M}\{\xi > t\}} = \frac{1 - \Phi(x)}{1 - \Phi(t)} \geq \frac{1 - (1 + \Phi(t))/2}{1 - \Phi(t)} = 0.5$$

and

$$\frac{\mathcal{M}\{(\xi \leq x) \cap (\xi > t)\}}{\mathcal{M}\{\xi > t\}} \leq \frac{\Phi(x)}{1 - \Phi(t)}.$$

It follows from the maximum uncertainty principle that

$$\Phi(x|(t, +\infty)) = \frac{\Phi(x)}{1 - \Phi(t)} \wedge 0.5.$$

When $(1 + \Phi(t))/2 \leq \Phi(x)$, we have $x \geq t$, and

$$\frac{\mathcal{M}\{(\xi > x) \cap (\xi > t)\}}{\mathcal{M}\{\xi > t\}} = \frac{1 - \Phi(x)}{1 - \Phi(t)} \leq \frac{1 - (1 + \Phi(t))/2}{1 - \Phi(t)} \leq 0.5.$$

Thus

$$\Phi(x|(t, +\infty)) = 1 - \frac{\mathcal{M}\{(\xi > x) \cap (\xi > t)\}}{\mathcal{M}\{\xi > t\}} = 1 - \frac{1 - \Phi(x)}{1 - \Phi(t)} = \frac{\Phi(x) - \Phi(t)}{1 - \Phi(t)}.$$

The theorem is proved.

Example 1.72: Let ξ be a linear uncertain variable $\mathcal{L}(a, b)$, and t a real number with $a < t < b$. Then the conditional uncertainty distribution of ξ given $\xi > t$ is

$$\Phi(x|(t, +\infty)) = \begin{cases} 0, & \text{if } x \leq t \\[2mm] \dfrac{x - a}{b - t} \wedge 0.5, & \text{if } t < x \leq (b + t)/2 \\[2mm] \dfrac{x - t}{b - t} \wedge 1, & \text{if } (b + t)/2 \leq x. \end{cases}$$

Theorem 1.71 *Let ξ be an uncertain variable with regular uncertainty distribution $\Phi(x)$, and t a real number with $\Phi(t) > 0$. Then the conditional*

uncertainty distribution of ξ given $\xi \leq t$ is

$$\Phi(x|(-\infty,t]) = \begin{cases} \dfrac{\Phi(x)}{\Phi(t)}, & \text{if } \Phi(x) \leq \Phi(t)/2 \\[2mm] \dfrac{\Phi(x) + \Phi(t) - 1}{\Phi(t)} \vee 0.5, & \text{if } \Phi(t)/2 \leq \Phi(x) < \Phi(t) \\[2mm] 1, & \text{if } \Phi(t) \leq \Phi(x). \end{cases}$$

Proof: It follows from $\Phi(x|(-\infty,t]) = \mathcal{M}\{\xi \leq x | \xi \leq t\}$ and the definition of conditional uncertainty that

$$\Phi(x|(-\infty,t]) = \begin{cases} \dfrac{\mathcal{M}\{(\xi \leq x) \cap (\xi \leq t)\}}{\mathcal{M}\{\xi \leq t\}}, & \text{if } \dfrac{\mathcal{M}\{(\xi \leq x) \cap (\xi \leq t)\}}{\mathcal{M}\{\xi \leq t\}} < 0.5 \\[3mm] 1 - \dfrac{\mathcal{M}\{(\xi > x) \cap (\xi \leq t)\}}{\mathcal{M}\{\xi \leq t\}}, & \text{if } \dfrac{\mathcal{M}\{(\xi > x) \cap (\xi \leq t)\}}{\mathcal{M}\{\xi \leq t\}} < 0.5 \\[3mm] 0.5, & \text{otherwise.} \end{cases}$$

When $\Phi(x) \leq \Phi(t)/2$, we have $x < t$, and

$$\frac{\mathcal{M}\{(\xi \leq x) \cap (\xi \leq t)\}}{\mathcal{M}\{\xi \leq t\}} = \frac{\Phi(x)}{\Phi(t)} \leq \frac{\Phi(t)/2}{\Phi(t)} = 0.5.$$

Thus

$$\Phi(x|(-\infty,t]) = \frac{\mathcal{M}\{(\xi \leq x) \cap (\xi \leq t)\}}{\mathcal{M}\{\xi \leq t\}} = \frac{\Phi(x)}{\Phi(t)}.$$

When $\Phi(t)/2 \leq \Phi(x) < \Phi(t)$, we have $x < t$, and

$$\frac{\mathcal{M}\{(\xi \leq x) \cap (\xi \leq t)\}}{\mathcal{M}\{\xi \leq t\}} = \frac{\Phi(x)}{\Phi(t)} \geq \frac{\Phi(t)/2}{\Phi(t)} = 0.5$$

and

$$\frac{\mathcal{M}\{(\xi > x) \cap (\xi \leq t)\}}{\mathcal{M}\{\xi \leq t\}} \leq \frac{1 - \Phi(x)}{\Phi(t)},$$

i.e.,

$$1 - \frac{\mathcal{M}\{(\xi > x) \cap (\xi \leq t)\}}{\mathcal{M}\{\xi \leq t\}} \geq \frac{\Phi(x) + \Phi(t) - 1}{\Phi(t)}.$$

It follows from the maximum uncertainty principle that

$$\Phi(x|(-\infty,t]) = \frac{\Phi(x) + \Phi(t) - 1}{\Phi(t)} \vee 0.5.$$

When $\Phi(t) \leq \Phi(x)$, we have $x \geq t$, and

$$\frac{\mathcal{M}\{(\xi > x) \cap (\xi \leq t)\}}{\mathcal{M}\{\xi \leq t\}} = \frac{\mathcal{M}\{\emptyset\}}{\Phi(t)} = 0 < 0.5.$$

Thus

$$\Phi(x|(-\infty, t]) = 1 - \frac{\mathcal{M}\{(\xi > x) \cap (\xi \leq t)\}}{\mathcal{M}\{\xi \leq t\}} = 1 - 0 = 1.$$

The theorem is proved.

Example 1.73: Let ξ be a linear uncertain variable $\mathcal{L}(a, b)$, and t a real number with $a < t < b$. Then the conditional uncertainty distribution of ξ given $\xi \leq t$ is

$$\Phi(x|(-\infty, t]) = \begin{cases} \dfrac{x-a}{t-a} \vee 0, & \text{if } x \leq (a+t)/2 \\[2mm] \left(1 - \dfrac{b-x}{t-a}\right) \vee 0.5, & \text{if } (a+t)/2 \leq x < t \\[2mm] 1, & \text{if } x \leq t. \end{cases}$$

Definition 1.36 *(Liu [127]) Let ξ be an uncertain variable. Then the conditional expected value of ξ given B is defined by*

$$E[\xi|B] = \int_0^{+\infty} \mathcal{M}\{\xi \geq r|B\}\mathrm{d}r - \int_{-\infty}^0 \mathcal{M}\{\xi \leq r|B\}\mathrm{d}r \qquad (1.225)$$

provided that at least one of the two integrals is finite.

Chapter 2

Uncertain Programming

Uncertain programming was founded by Liu [129] in 2009 as a type of mathematical programming involving uncertain variables. This chapter provides a general framework of uncertain programming, including expected value model, chance-constrained programming, dependent-chance programming, uncertain dynamic programming and uncertain multilevel programming. Finally, we present some uncertain programming models for system reliability design, project scheduling problem, vehicle routing problem, facility location problem, and machine scheduling problem.

2.1 Ranking Criteria

Assume that $\boldsymbol{x}$ is a decision vector, $\boldsymbol{\xi}$ is an uncertain vector, $f(\boldsymbol{x}, \boldsymbol{\xi})$ is a return function, and $g_j(\boldsymbol{x}, \boldsymbol{\xi})$ are constraint functions, $j = 1, 2, \cdots, p$. Let us examine

$$\begin{cases} \max f(\boldsymbol{x}, \boldsymbol{\xi}) \\ \text{subject to:} \\ \quad g_j(\boldsymbol{x}, \boldsymbol{\xi}) \leq 0, \quad j = 1, 2, \cdots, p. \end{cases} \tag{2.1}$$

The model (2.1) is only a conceptual model rather than a mathematical model because we cannot maximize an uncertain quantity just like that we cannot maximize a random quantity. In other words, there does not exist a natural ordership in an uncertain world.

Thus an important problem appearing in this area is how to rank uncertain variables. Let ξ and η be two uncertain variables. Here we give four ranking criteria.

Expected Value Criterion: We say $\xi > \eta$ if and only if $E[\xi] > E[\eta]$.

Optimistic Value Criterion: We say $\xi > \eta$ if and only if, for some predetermined confidence level $\alpha \in (0, 1]$, we have $\xi_{\sup}(\alpha) > \eta_{\sup}(\alpha)$, where $\xi_{\sup}(\alpha)$ and $\eta_{\sup}(\alpha)$ are the α-optimistic values of ξ and η, respectively.

Pessimistic Value Criterion: We say $\xi > \eta$ if and only if, for some predetermined confidence level $\alpha \in (0, 1]$, we have $\xi_{\inf}(\alpha) > \eta_{\inf}(\alpha)$, where $\xi_{\inf}(\alpha)$ and $\eta_{\inf}(\alpha)$ are the α-pessimistic values of ξ and η, respectively.

Chance Criterion: We say $\xi > \eta$ if and only if, for some predetermined levels $\bar{r}$, we have $\mathcal{M}\{\xi \geq \bar{r}\} > \mathcal{M}\{\eta \geq \bar{r}\}$.

2.2   Expected Value Model

Assume that we believe the expected value criterion. In order to obtain a decision with maximum expected return subject to expected constraints, we have the following expected value model,

$$\begin{cases} \max E[f(\boldsymbol{x}, \boldsymbol{\xi})] \\ \text{subject to:} \\ \quad E[g_j(\boldsymbol{x}, \boldsymbol{\xi})] \leq 0, \ j = 1, 2, \cdots, p \end{cases} \tag{2.2}$$

where $\boldsymbol{x}$ is a decision vector, $\boldsymbol{\xi}$ is an uncertain vector, f is a return function, and g_j are constraint functions for $j = 1, 2, \cdots, p$.

Definition 2.1 *A solution x is feasible if and only if $E[g_j(\boldsymbol{x}, \boldsymbol{\xi})] \leq 0$ for $j = 1, 2, \cdots, p$. A feasible solution $\boldsymbol{x}^*$ is an optimal solution to the expected value model (2.2) if $E[f(\boldsymbol{x}^*, \boldsymbol{\xi})] \geq E[f(\boldsymbol{x}, \boldsymbol{\xi})]$ for any feasible solution $\boldsymbol{x}$.*

In practice, a decision maker may want to optimize multiple objectives. Thus we have the following expected value multiobjective programming,

$$\begin{cases} \max [E[f_1(\boldsymbol{x}, \boldsymbol{\xi})], E[f_2(\boldsymbol{x}, \boldsymbol{\xi})], \cdots, E[f_m(\boldsymbol{x}, \boldsymbol{\xi})]] \\ \text{subject to:} \\ \quad E[g_j(\boldsymbol{x}, \boldsymbol{\xi})] \leq 0, \ j = 1, 2, \cdots, p \end{cases} \tag{2.3}$$

where $f_i(\boldsymbol{x}, \boldsymbol{\xi})$ are return functions for $i = 1, 2, \cdots, m$, and $g_j(\boldsymbol{x}, \boldsymbol{\xi})$ are constraint functions for $j = 1, 2, \cdots, p$.

Definition 2.2 *A feasible solution $\boldsymbol{x}^*$ is said to be a Pareto solution to the expected value multiobjective programming (2.3) if there is no feasible solution $\boldsymbol{x}$ such that*

$$E[f_i(\boldsymbol{x}, \boldsymbol{\xi})] \geq E[f_i(\boldsymbol{x}^*, \boldsymbol{\xi})], \quad i = 1, 2, \cdots, m \tag{2.4}$$

and $E[f_j(\boldsymbol{x}, \boldsymbol{\xi})] > E[f_j(\boldsymbol{x}^*, \boldsymbol{\xi})]$ for at least one index j.

In order to balance multiple conflicting objectives, a decision-maker may establish a hierarchy of importance among these incompatible goals so as to satisfy as many goals as possible in the order specified. Thus we have an

expected value goal programming,

$$
\begin{cases}
\min \sum_{j=1}^{l} P_j \sum_{i=1}^{m} (u_{ij} d_i^+ \vee 0 + v_{ij} d_i^- \vee 0) \\
\text{subject to:} \\
\quad E[f_i(\boldsymbol{x}, \boldsymbol{\xi})] - b_i = d_i^+, \quad i = 1, 2, \cdots, m \\
\quad b_i - E[f_i(\boldsymbol{x}, \boldsymbol{\xi})] = d_i^-, \quad i = 1, 2, \cdots, m \\
\quad E[g_j(\boldsymbol{x}, \boldsymbol{\xi})] \leq 0, \qquad\quad j = 1, 2, \cdots, p
\end{cases}
\tag{2.5}
$$

where P_j is the preemptive priority factor which expresses the relative importance of various goals, $P_j \gg P_{j+1}$, for all j, u_{ij} is the weighting factor corresponding to positive deviation for goal i with priority j assigned, v_{ij} is the weighting factor corresponding to negative deviation for goal i with priority j assigned, $d_i^+ \vee 0$ is the positive deviation from the target of goal i, $d_i^- \vee 0$ is the negative deviation from the target of goal i, f_i is a function in goal constraints, g_j is a function in real constraints, b_i is the target value according to goal i, l is the number of priorities, m is the number of goal constraints, and p is the number of real constraints.

Theorem 2.1 *Assume $f(\boldsymbol{x}, \boldsymbol{\xi}) = h_1(\boldsymbol{x})\xi_1 + h_2(\boldsymbol{x})\xi_2 + \cdots + h_n(\boldsymbol{x})\xi_n + h_0(\boldsymbol{x})$ where $h_1(\boldsymbol{x}), h_2(\boldsymbol{x}), \cdots, h_n(\boldsymbol{x}), h_0(\boldsymbol{x})$ are real-valued functions and $\xi_1, \xi_2, \cdots, \xi_n$ are independent uncertain variables. Then*

$$
E[f(\boldsymbol{x}, \boldsymbol{\xi})] = h_1(\boldsymbol{x})E[\xi_1] + h_2(\boldsymbol{x})E[\xi_2] + \cdots + h_n(\boldsymbol{x})E[\xi_n] + h_0(\boldsymbol{x}).
\tag{2.6}
$$

Proof: It follows from the linearity of expected value operator immediately.

Theorem 2.2 *Assume that $\xi_1, \xi_2, \cdots, \xi_n$ are independent uncertain variables and $h_1(\boldsymbol{x}), h_2(\boldsymbol{x}), \cdots, h_n(\boldsymbol{x}), h_0(\boldsymbol{x})$ are real-valued functions. Then*

$$
E[h_1(\boldsymbol{x})\xi_1 + h_2(\boldsymbol{x})\xi_2 + \cdots + h_n(\boldsymbol{x})\xi_n + h_0(\boldsymbol{x})] \leq 0
\tag{2.7}
$$

holds if and only if

$$
h_1(\boldsymbol{x})E[\xi_1] + h_2(\boldsymbol{x})E[\xi_2] + \cdots + h_n(\boldsymbol{x})E[\xi_n] + h_0(\boldsymbol{x}) \leq 0.
\tag{2.8}
$$

Proof: It follows from Theorem 2.1 immediately.

2.3   Chance-Constrained Programming

Since the uncertain constraints $g_j(\boldsymbol{x}, \boldsymbol{\xi}) \leq 0, j = 1, 2, \cdots, p$ do not define a deterministic feasible set, it is naturally desired that the uncertain constraints hold with a confidence level α. Then we have a chance constraint as follows,

$$
\mathcal{M}\{g_j(\boldsymbol{x}, \boldsymbol{\xi}) \leq 0, j = 1, 2, \cdots, p\} \geq \alpha.
\tag{2.9}
$$

Maximax Chance-Constrained Programming

Assume that we believe the optimistic value criterion. If we want to maximize the optimistic value to the uncertain return subject to some chance constraints, then we have the following maximax chance-constrained programming,

$$
\begin{cases}
\max\limits_{\boldsymbol{x}} \max\limits_{\overline{f}} \overline{f} \\
\text{subject to:} \\
\quad \mathcal{M}\left\{ f(\boldsymbol{x},\boldsymbol{\xi}) \geq \overline{f} \right\} \geq \beta \\
\quad \mathcal{M}\left\{ g_j(\boldsymbol{x},\boldsymbol{\xi}) \leq 0 \right\} \geq \alpha_j, \quad j = 1, 2, \cdots, p
\end{cases}
\tag{2.10}
$$

where α_j and β are specified confidence levels for $j = 1, 2, \cdots, p$, and $\max \overline{f}$ is the β-optimistic return.

In practice, it is possible that there exist multiple objectives. We thus have the following maximax chance-constrained multiobjective programming,

$$
\begin{cases}
\max\limits_{\boldsymbol{x}} \left[\max\limits_{\overline{f}_1} \overline{f}_1, \max\limits_{\overline{f}_2} \overline{f}_2, \cdots, \max\limits_{\overline{f}_m} \overline{f}_m \right] \\
\text{subject to:} \\
\quad \mathcal{M}\left\{ f_i(\boldsymbol{x},\boldsymbol{\xi}) \geq \overline{f}_i \right\} \geq \beta_i, \quad i = 1, 2, \cdots, m \\
\quad \mathcal{M}\left\{ g_j(\boldsymbol{x},\boldsymbol{\xi}) \leq 0 \right\} \geq \alpha_j, \quad j = 1, 2, \cdots, p
\end{cases}
\tag{2.11}
$$

where β_i are predetermined confidence levels for $i = 1, 2, \cdots, m$, and $\max \overline{f}_i$ are the β-optimistic values to the return functions $f_i(\boldsymbol{x},\boldsymbol{\xi})$, $i = 1, 2, \cdots, m$, respectively.

If the priority structure and target levels are set by the decision-maker, then we have a minimin chance-constrained goal programming,

$$
\begin{cases}
\min\limits_{\boldsymbol{x}} \sum\limits_{j=1}^{l} P_j \sum\limits_{i=1}^{m} \left(u_{ij} \left(\min\limits_{d_i^+} d_i^+ \vee 0 \right) + v_{ij} \left(\min\limits_{d_i^-} d_i^- \vee 0 \right) \right) \\
\text{subject to:} \\
\quad \mathcal{M}\left\{ f_i(\boldsymbol{x},\boldsymbol{\xi}) - b_i \leq d_i^+ \right\} \geq \beta_i^+, \quad i = 1, 2, \cdots, m \\
\quad \mathcal{M}\left\{ b_i - f_i(\boldsymbol{x},\boldsymbol{\xi}) \leq d_i^- \right\} \geq \beta_i^-, \quad i = 1, 2, \cdots, m \\
\quad \mathcal{M}\left\{ g_j(\boldsymbol{x},\boldsymbol{\xi}) \leq 0 \right\} \geq \alpha_j, \qquad\quad j = 1, 2, \cdots, p
\end{cases}
\tag{2.12}
$$

where P_j is the preemptive priority factor which expresses the relative importance of various goals, $P_j \gg P_{j+1}$, for all j, u_{ij} is the weighting factor corresponding to positive deviation for goal i with priority j assigned, v_{ij} is the weighting factor corresponding to negative deviation for goal i with priority j assigned, $\min d_i^+ \vee 0$ is the β_i^+-optimistic positive deviation from the target of goal i, $\min d_i^- \vee 0$ is the β_i^--optimistic negative deviation from the target of goal i, b_i is the target value according to goal i, and l is the number of priorities.

Minimax Chance-Constrained Programming

Assume that we believe the pessimistic value criterion. If we want to maximize the pessimistic value subject to some chance constraints, then we have the following minimax chance-constrained programming,

$$
\begin{cases}
\max\limits_{\boldsymbol{x}} \min\limits_{\overline{f}} \overline{f} \\
\text{subject to:} \\
\quad \mathcal{M}\left\{f(\boldsymbol{x}, \boldsymbol{\xi}) \leq \overline{f}\right\} \geq \beta \\
\quad \mathcal{M}\left\{g_j(\boldsymbol{x}, \boldsymbol{\xi}) \leq 0\right\} \geq \alpha_j, \quad j = 1, 2, \cdots, p
\end{cases}
\tag{2.13}
$$

where α_j and β are specified confidence levels for $j = 1, 2, \cdots, p$, and $\min \overline{f}$ is the β-pessimistic return.

If there are multiple objectives, then we have the following minimax chance-constrained multiobjective programming,

$$
\begin{cases}
\max\limits_{\boldsymbol{x}} \left[\min\limits_{\overline{f}_1} \overline{f}_1, \ \min\limits_{\overline{f}_2} \overline{f}_2, \ \cdots, \ \min\limits_{\overline{f}_m} \overline{f}_m\right] \\
\text{subject to:} \\
\quad \mathcal{M}\left\{f_i(\boldsymbol{x}, \boldsymbol{\xi}) \leq \overline{f}_i\right\} \geq \beta_i, \quad i = 1, 2, \cdots, m \\
\quad \mathcal{M}\left\{g_j(\boldsymbol{x}, \boldsymbol{\xi}) \leq 0\right\} \geq \alpha_j, \quad j = 1, 2, \cdots, p
\end{cases}
\tag{2.14}
$$

where $\min \overline{f}_i$ are the β_i-pessimistic values to the return functions $f_i(\boldsymbol{x}, \boldsymbol{\xi})$, $i = 1, 2, \cdots, m$, respectively.

We can also formulate an uncertain decision system as a minimax chance-constrained goal programming according to the priority structure and target levels set by the decision-maker:

$$
\begin{cases}
\min\limits_{\boldsymbol{x}} \sum\limits_{j=1}^{l} P_j \sum\limits_{i=1}^{m} \left[u_{ij}\left(\max\limits_{d_i^+} d_i^+ \vee 0\right) + v_{ij}\left(\max\limits_{d_i^-} d_i^- \vee 0\right)\right] \\
\text{subject to:} \\
\quad \mathcal{M}\left\{f_i(\boldsymbol{x}, \boldsymbol{\xi}) - b_i \geq d_i^+\right\} \geq \beta_i^+, \quad i = 1, 2, \cdots, m \\
\quad \mathcal{M}\left\{b_i - f_i(\boldsymbol{x}, \boldsymbol{\xi}) \geq d_i^-\right\} \geq \beta_i^-, \quad i = 1, 2, \cdots, m \\
\quad \mathcal{M}\left\{g_j(\boldsymbol{x}, \boldsymbol{\xi}) \leq 0\right\} \geq \alpha_j, \qquad\quad j = 1, 2, \cdots, p
\end{cases}
\tag{2.15}
$$

where P_j is the preemptive priority factor which expresses the relative importance of various goals, $P_j \gg P_{j+1}$, for all j, u_{ij} is the weighting factor corresponding to positive deviation for goal i with priority j assigned, v_{ij} is the weighting factor corresponding to negative deviation for goal i with priority j assigned, $\max d_i^+ \vee 0$ is the β_i^+-pessimistic positive deviation from the target of goal i, $\max d_i^- \vee 0$ is the β_i^--pessimistic negative deviation from the target of goal i, b_i is the target value according to goal i, and l is the number of priorities.

Theorem 2.3 *Assume that $\xi_1, \xi_2, \cdots, \xi_n$ are independent uncertain variables with regular uncertainty distributions $\Phi_1, \Phi_2, \cdots, \Phi_n$, respectively, and $h_1(\boldsymbol{x}), h_2(\boldsymbol{x}), \cdots, h_n(\boldsymbol{x}), h_0(\boldsymbol{x})$ are real-valued functions. Then*

$$\mathcal{M}\left\{\sum_{i=1}^{n} h_i(\boldsymbol{x})\xi_i \leq h_0(\boldsymbol{x})\right\} \geq \alpha \tag{2.16}$$

holds if and only if

$$\sum_{i=1}^{n} h_i^+(\boldsymbol{x})\Phi_i^{-1}(\alpha) - \sum_{i=1}^{n} h_i^-(\boldsymbol{x})\Phi_i^{-1}(1-\alpha) \leq h_0(\boldsymbol{x}) \tag{2.17}$$

where

$$h_i^+(\boldsymbol{x}) = \begin{cases} h_i(\boldsymbol{x}), & \text{if } h_i(\boldsymbol{x}) > 0 \\ 0, & \text{if } h_i(\boldsymbol{x}) \leq 0, \end{cases} \tag{2.18}$$

$$h_i^-(\boldsymbol{x}) = \begin{cases} 0, & \text{if } h_i(\boldsymbol{x}) \geq 0 \\ -h_i(\boldsymbol{x}), & \text{if } h_i(\boldsymbol{x}) < 0 \end{cases} \tag{2.19}$$

for $i = 1, 2, \cdots, n$. Especially, if $h_1(\boldsymbol{x}), h_2(\boldsymbol{x}), \cdots, h_n(\boldsymbol{x})$ are all nonnegative, then (2.17) becomes

$$\sum_{i=1}^{n} h_i(\boldsymbol{x})\Phi_i^{-1}(\alpha) \leq h_0(\boldsymbol{x}); \tag{2.20}$$

if $h_1(\boldsymbol{x}), h_2(\boldsymbol{x}), \cdots, h_n(\boldsymbol{x})$ are all nonpositive, then (2.17) becomes

$$\sum_{i=1}^{n} h_i(\boldsymbol{x})\Phi_i^{-1}(1-\alpha) \leq h_0(\boldsymbol{x}). \tag{2.21}$$

Proof: For each i, if $h_i(x) > 0$, then $h_i(\boldsymbol{x})\xi_i$ is an uncertain variable whose uncertainty distribution is described by

$$\Psi_i^{-1}(\alpha) = h_i^+(\boldsymbol{x})\Phi_i^{-1}(\alpha), \quad 0 < \alpha < 1.$$

If $h_i(x) < 0$, then $h_i(\boldsymbol{x})\xi_i$ is an uncertain variable whose uncertainty distribution is described by

$$\Psi_i^{-1}(\alpha) = -h_i^-(\boldsymbol{x})\Phi_i^{-1}(1-\alpha), \quad 0 < \alpha < 1.$$

It follows from the operational law that the uncertainty distribution of the sum $h_1(\boldsymbol{x})\xi_1 + h_2(\boldsymbol{x})\xi_2 + \cdots + h_n(\boldsymbol{x})\xi_n$ is described by

$$\Psi^{-1}(\alpha) = \Psi_1^{-1}(\alpha) + \Psi_2^{-1}(\alpha) + \cdots + \Psi_n^{-1}(\alpha), \quad 0 < \alpha < 1.$$

From which we may derive the result immediately.

Theorem 2.4 *Assume that $x_1, x_2, \cdots, x_n$ are nonnegative decision variables, and $\xi_1, \xi_2, \cdots, \xi_n, \xi$ are independently linear uncertain variables $\mathcal{L}(a_1, b_1)$, $\mathcal{L}(a_2, b_2), \cdots, \mathcal{L}(a_n, b_n), \mathcal{L}(a, b)$, respectively. Then for any confidence level $\alpha \in (0, 1)$, the chance constraint*

$$\mathcal{M}\left\{\sum_{i=1}^{n} \xi_i x_i \leq \xi\right\} \geq \alpha \qquad (2.22)$$

holds if and only if

$$\sum_{i=1}^{n}((1 - \alpha)a_i + \alpha b_i)x_i \leq \alpha a + (1 - \alpha)b. \qquad (2.23)$$

Proof: Assume that the uncertain variables $\xi_1, \xi_2, \cdots, \xi_n, \xi$ have uncertainty distributions $\Phi_1, \Phi_2, \cdots, \Phi_n, \Phi$, respectively. Then

$$\Phi_i^{-1}(\alpha) = (1 - \alpha)a_i + \alpha b_i, \quad i = 1, 2, \cdots, n,$$

$$\Phi^{-1}(1 - \alpha) = \alpha a + (1 - \alpha)b.$$

Thus the result follows from Theorem 2.3 immediately.

Theorem 2.5 *Assume that $x_1, x_2, \cdots, x_n$ are nonnegative decision variables, and $\xi_1, \xi_2, \cdots, \xi_n, \xi$ are independently zigzag uncertain variables $\mathcal{Z}(a_1, b_1, c_1)$, $\mathcal{Z}(a_2, b_2, c_2), \cdots, \mathcal{Z}(a_n, b_n, c_n), \mathcal{Z}(a, b, c)$, respectively. Then for any confidence level $\alpha \geq 0.5$, the chance constraint*

$$\mathcal{M}\left\{\sum_{i=1}^{n} \xi_i x_i \leq \xi\right\} \geq \alpha \qquad (2.24)$$

holds if and only if

$$\sum_{i=1}^{n}((2 - 2\alpha)b_i + (2\alpha - 1)c_i)x_i \leq \alpha(2\alpha - 1)a + (2 - 2\alpha)b. \qquad (2.25)$$

Proof: Assume that the uncertain variables $\xi_1, \xi_2, \cdots, \xi_n, \xi$ have uncertainty distributions $\Phi_1, \Phi_2, \cdots, \Phi_n, \Phi$, respectively. Then

$$\Phi_i^{-1}(\alpha) = (2 - 2\alpha)b_i + (2\alpha - 1)c_i, \quad i = 1, 2, \cdots, n,$$

$$\Phi^{-1}(1 - \alpha) = (2\alpha - 1)a + (2 - 2\alpha)b.$$

Thus the result follows from Theorem 2.3 immediately.

Theorem 2.6 *Assume that $x_1, x_2, \cdots, x_n$ are nonnegative decision variables, and $\xi_1, \xi_2, \cdots, \xi_n, \xi$ are independently normal uncertain variables $\mathcal{N}(e_1, \sigma_1)$,*

$\mathcal{N}(e_2, \sigma_2), \cdots, \mathcal{N}(e_n, \sigma_n), \mathcal{N}(e, \sigma)$, respectively. Then for any confidence level $\alpha \in (0, 1)$, the chance constraint

$$\mathcal{M}\left\{\sum_{i=1}^{n} \xi_i x_i \leq \xi\right\} \geq \alpha \qquad (2.26)$$

holds if and only if

$$\sum_{i=1}^{n} \left(e_i + \frac{\sigma_i\sqrt{3}}{\pi} \ln \frac{\alpha}{1-\alpha}\right) x_i \leq e - \frac{\sigma\sqrt{3}}{\pi} \ln \frac{\alpha}{1-\alpha}. \qquad (2.27)$$

Proof: Assume that the uncertain variables $\xi_1, \xi_2, \cdots, \xi_n, \xi$ have uncertainty distributions $\Phi_1, \Phi_2, \cdots, \Phi_n, \Phi$, respectively. Then

$$\Phi_i^{-1}(\alpha) = e_i + \frac{\sigma_i\sqrt{3}}{\pi} \ln \frac{\alpha}{1-\alpha}, \quad i = 1, 2, \cdots, n,$$

$$\Phi^{-1}(1-\alpha) = e - \frac{\sigma\sqrt{3}}{\pi} \ln \frac{\alpha}{1-\alpha}.$$

Thus the result follows from Theorem 2.3 immediately.

Theorem 2.7 *Assume $x_1, x_2, \cdots, x_n$ are nonnegative decision variables, and $\xi_1, \xi_2, \cdots, \xi_n, \xi$ are independently lognormal uncertain variables $\mathcal{LOGN}(e_1, \sigma_1)$, $\mathcal{LOGN}(e_2, \sigma_2), \cdots, \mathcal{LOGN}(e_n, \sigma_n), \mathcal{LOGN}(e, \sigma)$, respectively. Then for any confidence level $\alpha \in (0, 1)$, the chance constraint*

$$\mathcal{M}\left\{\sum_{i=1}^{n} \xi_i x_i \leq \xi\right\} \geq \alpha \qquad (2.28)$$

holds if and only if

$$\sum_{i=1}^{n} \exp(e_i) \left(\frac{\alpha}{1-\alpha}\right)^{\sqrt{3}\sigma_i/\pi} x_i \leq \exp(e) \left(\frac{1-\alpha}{\alpha}\right)^{\sqrt{3}\sigma/\pi}. \qquad (2.29)$$

Proof: Assume that the uncertain variables $\xi_1, \xi_2, \cdots, \xi_n, \xi$ have uncertainty distributions $\Phi_1, \Phi_2, \cdots, \Phi_n, \Phi$, respectively. Then

$$\Phi_i^{-1}(\alpha) = \exp(e_i) \left(\frac{\alpha}{1-\alpha}\right)^{\sqrt{3}\sigma_i/\pi}, \quad i = 1, 2, \cdots, n,$$

$$\Phi^{-1}(1-\alpha) = \exp(e) \left(\frac{1-\alpha}{\alpha}\right)^{\sqrt{3}\sigma/\pi}.$$

Thus the result follows from Theorem 2.3 immediately.

2.4 Dependent-Chance Programming

In practice, there usually exist multiple tasks in a complex uncertain decision system. Sometimes, the decision-maker believes the chance criterion and wishes to maximize the chance of meeting these tasks. In order to model this type of uncertain decision system, Liu [129] provided the third type of uncertain programming, called *dependent-chance programming*, in which the underlying philosophy is based on selecting the decision with maximal chance to meet the task. Dependent-chance programming breaks the concept of feasible set and replaces it with uncertain environment.

Definition 2.3 *By an uncertain environment we mean the uncertain constraints represented by*

$$g_j(\boldsymbol{x}, \boldsymbol{\xi}) \le 0, \quad j = 1, 2, \cdots, p \tag{2.30}$$

where $\boldsymbol{x}$ is a decision vector, and $\boldsymbol{\xi}$ is an uncertain vector.

Definition 2.4 *By a task we mean an uncertain inequality (or a system of uncertain inequalities) represented by*

$$h(\boldsymbol{x}, \boldsymbol{\xi}) \le 0 \tag{2.31}$$

where $\boldsymbol{x}$ is a decision vector, and $\boldsymbol{\xi}$ is an uncertain vector.

Definition 2.5 *The chance function of task $\mathcal{E}$ characterized by (2.31) is defined as the uncertain measure that the task $\mathcal{E}$ is met, i.e.,*

$$f(\boldsymbol{x}) = \mathcal{M}\{h(\boldsymbol{x}, \boldsymbol{\xi}) \le 0\} \tag{2.32}$$

subject to the uncertain environment (2.30).

How do we compute the chance function in an uncertain environment? In order to answer this question, we first give some basic definitions. Let $r(x_1, x_2, \cdots, x_n)$ be an n-dimensional function. The ith decision variable x_i is said to be degenerate if

$$r(x_1, \cdots, x_{i-1}, x_i', x_{i+1}, \cdots, x_n) = r(x_1, \cdots, x_{i-1}, x_i'', x_{i+1}, \cdots, x_n)$$

for any x_i' and x_i''; otherwise it is nondegenerate. For example,

$$r(x_1, x_2, x_3, x_4, x_5) = (x_1 + x_3)/x_4$$

is a 5-dimensional function. The variables x_1, x_3, x_4 are nondegenerate, but x_2 and x_5 are degenerate.

Definition 2.6 *Let $\mathcal{E}$ be a task $h(\boldsymbol{x}, \boldsymbol{\xi}) \le 0$. The support of the task $\mathcal{E}$, denoted by $\mathcal{E}^*$, is defined as the set consisting of all nondegenerate decision variables of $h(\boldsymbol{x}, \boldsymbol{\xi})$.*

Definition 2.7 *The jth constraint $g_j(\boldsymbol{x}, \boldsymbol{\xi}) \leq 0$ is called an active constraint of task $\mathcal{E}$ if the set of nondegenerate decision variables of $g_j(\boldsymbol{x}, \boldsymbol{\xi})$ and the support $\mathcal{E}^*$ have nonempty intersection; otherwise it is inactive.*

Definition 2.8 *Let $\mathcal{E}$ be a task $h(\boldsymbol{x}, \boldsymbol{\xi}) \leq 0$ in the uncertain environment $g_j(\boldsymbol{x}, \boldsymbol{\xi}) \leq 0$, $j = 1, 2, \cdots, p$. The dependent support of task $\mathcal{E}$, denoted by $\mathcal{E}^{**}$, is defined as the set consisting of all nondegenerate decision variables of $h(\boldsymbol{x}, \boldsymbol{\xi})$ and $g_j(\boldsymbol{x}, \boldsymbol{\xi})$ in the active constraints of task $\mathcal{E}$.*

Remark 2.1: It is obvious that $\mathcal{E}^* \subset \mathcal{E}^{**}$ holds.

Definition 2.9 *The jth constraint $g_j(\boldsymbol{x}, \boldsymbol{\xi}) \leq 0$ is called a dependent constraint of task $\mathcal{E}$ if the set of nondegenerate decision variables of $g_j(\boldsymbol{x}, \boldsymbol{\xi})$ and the dependent support $\mathcal{E}^{**}$ have nonempty intersection; otherwise it is independent.*

Remark 2.2: An active constraint must be a dependent constraint.

Definition 2.10 *Let $\mathcal{E}$ be a task $h(\boldsymbol{x}, \boldsymbol{\xi}) \leq 0$ in the uncertain environment $g_j(\boldsymbol{x}, \boldsymbol{\xi}) \leq 0$, $j = 1, 2, \cdots, p$. For each decision $\boldsymbol{x}$ and realization $\boldsymbol{\xi}$, the task $\mathcal{E}$ is said to be consistent in the uncertain environment if the following two conditions hold: (i) $h(\boldsymbol{x}, \boldsymbol{\xi}) \leq 0$; and (ii) $g_j(\boldsymbol{x}, \boldsymbol{\xi}) \leq 0$, $j \in J$, where J is the index set of all dependent constraints.*

In order to maximize the chance of some task in an uncertain environment, a dependent-chance programming may be formulated as follows,

$$\begin{cases} \max \mathcal{M}\{h(\boldsymbol{x}, \boldsymbol{\xi}) \leq 0\} \\ \text{subject to:} \\ \quad g_j(\boldsymbol{x}, \boldsymbol{\xi}) \leq 0, \quad j = 1, 2, \cdots, p \end{cases} \tag{2.33}$$

where $\boldsymbol{x}$ is an n-dimensional decision vector, $\boldsymbol{\xi}$ is an uncertain vector, the task $\mathcal{E}$ is characterized by $h(\boldsymbol{x}, \boldsymbol{\xi}) \leq 0$, and the uncertain environment is described by the uncertain constraints $g_j(\boldsymbol{x}, \boldsymbol{\xi}) \leq 0$, $j = 1, 2, \cdots, p$. The model (2.33) is equivalent to

$$\max \mathcal{M}\{h(\boldsymbol{x}, \boldsymbol{\xi}) \leq 0, \, g_j(\boldsymbol{x}, \boldsymbol{\xi}) \leq 0, j \in J\} \tag{2.34}$$

where J is the index set of all dependent constraints.

If there are multiple tasks in an uncertain environment, then we have the following dependent-chance multiobjective programming,

$$\begin{cases} \max [\mathcal{M}\{h_1(\boldsymbol{x}, \boldsymbol{\xi}) \leq 0\}, \cdots, \mathcal{M}\{h_m(\boldsymbol{x}, \boldsymbol{\xi}) \leq 0\}] \\ \text{subject to:} \\ \quad g_j(\boldsymbol{x}, \boldsymbol{\xi}) \leq 0, \quad j = 1, 2, \cdots, p \end{cases} \tag{2.35}$$

where tasks $\mathcal{E}_i$ are characterized by $h_i(\boldsymbol{x}, \boldsymbol{\xi}) \leq 0$, $i = 1, 2, \cdots, m$, respectively. The model (2.35) is equivalent to

$$
\begin{cases}
\max \mathcal{M}\{h_1(\boldsymbol{x}, \boldsymbol{\xi}) \leq 0,\ g_j(\boldsymbol{x}, \boldsymbol{\xi}) \leq 0, j \in J_1\} \\
\max \mathcal{M}\{h_2(\boldsymbol{x}, \boldsymbol{\xi}) \leq 0,\ g_j(\boldsymbol{x}, \boldsymbol{\xi}) \leq 0, j \in J_2\} \\
\quad \cdots \\
\max \mathcal{M}\{h_m(\boldsymbol{x}, \boldsymbol{\xi}) \leq 0,\ g_j(\boldsymbol{x}, \boldsymbol{\xi}) \leq 0, j \in J_m\}
\end{cases}
\tag{2.36}
$$

where J_i are the index sets of all dependent constraints of tasks $\mathcal{E}_i$, $i = 1, 2, \cdots, m$, respectively.

Dependent-chance goal programming is employed to formulate uncertain decision systems according to the priority structure and target levels set by the decision-maker,

$$
\begin{cases}
\min \sum\limits_{j=1}^{l} P_j \sum\limits_{i=1}^{m} (u_{ij} d_i^+ \vee 0 + v_{ij} d_i^- \vee 0) \\
\text{subject to:} \\
\quad \mathcal{M}\{h_i(\boldsymbol{x}, \boldsymbol{\xi}) \leq 0\} - b_i = d_i^+, \quad i = 1, 2, \cdots, m \\
\quad b_i - \mathcal{M}\{h_i(\boldsymbol{x}, \boldsymbol{\xi}) \leq 0\} = d_i^-, \quad i = 1, 2, \cdots, m \\
\quad g_j(\boldsymbol{x}, \boldsymbol{\xi}) \leq 0, \quad\quad\quad\quad\quad\quad j = 1, 2, \cdots, p
\end{cases}
$$

where P_j is the preemptive priority factor which expresses the relative importance of various goals, $P_j \gg P_{j+1}$, for all j, u_{ij} is the weighting factor corresponding to positive deviation for goal i with priority j assigned, v_{ij} is the weighting factor corresponding to negative deviation for goal i with priority j assigned, $d_i^+ \vee 0$ is the positive deviation from the target of goal i, $d_i^- \vee 0$ is the negative deviation from the target of goal i, g_j is a function in system constraints, b_i is the target value according to goal i, l is the number of priorities, m is the number of goal constraints, and p is the number of system constraints.

Theorem 2.8 *Assume $x_1, x_2, \cdots, x_n$ are nonnegative decision variables, and $\xi_1, \xi_2, \cdots, \xi_n$ are independently linear uncertain variables $\mathcal{L}(a_1, b_1)$, $\mathcal{L}(a_2, b_2)$, $\cdots, \mathcal{L}(a_n, b_n)$, respectively. When*

$$
t \in \left[\sum_{i=1}^{n} a_i x_i, \sum_{i=1}^{n} b_i x_i \right],
\tag{2.37}
$$

we have

$$
\mathcal{M}\left\{ \sum_{i=1}^{n} \xi_i x_i \leq t \right\} = \frac{t - \sum\limits_{i=1}^{n} a_i x_i}{\sum\limits_{i=1}^{n} (b_i - a_i) x_i}.
\tag{2.38}
$$

Otherwise, the measure will be 0 if t is on the left-hand side of interval (2.37) or 1 if t is on the right-hand side.

Proof: Since $\xi_1, \xi_2, \cdots, \xi_n$ are independently linear uncertain variables, their weighted sum $\xi_1 x_1 + \xi_2 x_2 + \cdots + \xi_n x_n$ is also a linear uncertain variable

$$\mathcal{L}\left(\sum_{i=1}^{n} a_i x_i, \sum_{i=1}^{n} b_i x_i\right).$$

From this fact we may derive the result immediately.

Theorem 2.9 *Assume that $x_1, x_2, \cdots, x_n$ are nonnegative decision variables, and $\xi_1, \xi_2, \cdots, \xi_n$ are independently zigzag uncertain variables $\mathcal{Z}(a_1, b_1, c_1)$, $\mathcal{Z}(a_2, b_2, c_2), \cdots, \mathcal{Z}(a_n, b_n, c_n)$, respectively. When*

$$t \in \left[\sum_{i=1}^{n} a_i x_i, \sum_{i=1}^{n} b_i x_i\right], \qquad (2.39)$$

we have

$$\mathcal{M}\left\{\sum_{i=1}^{n} \xi_i x_i \leq t\right\} = \frac{t - \displaystyle\sum_{i=1}^{n} a_i x_i}{2 \displaystyle\sum_{i=1}^{n} (b_i - a_i) x_i}. \qquad (2.40)$$

When

$$t \in \left[\sum_{i=1}^{n} b_i x_i, \sum_{i=1}^{n} c_i x_i\right], \qquad (2.41)$$

we have

$$\mathcal{M}\left\{\sum_{i=1}^{n} \xi_i x_i \leq t\right\} = \frac{t + \displaystyle\sum_{i=1}^{n} (c_i - 2b_i) x_i}{2 \displaystyle\sum_{i=1}^{n} (c_i - b_i) x_i}. \qquad (2.42)$$

Otherwise, the measure will be 0 if t is on the left-hand side of interval (2.39) or 1 if t is on the right-hand side of interval (2.41).

Proof: Since $\xi_1, \xi_2, \cdots, \xi_n$ are independently zigzag uncertain variables, their weighted sum $\xi_1 x_1 + \xi_2 x_2 + \cdots + \xi_n x_n$ is also a zigzag uncertain variable

$$\mathcal{Z}\left(\sum_{i=1}^{n} a_i x_i, \sum_{i=1}^{n} b_i x_i, \sum_{i=1}^{n} c_i x_i\right).$$

From this fact we may derive the result immediately.

Theorem 2.10 *Assume that $x_1, x_2, \cdots, x_n$ are nonnegative decision variables, and $\xi_1, \xi_2, \cdots, \xi_n$ are independently normal uncertain variables $\mathcal{N}(e_1, \sigma_1)$, $\mathcal{N}(e_2, \sigma_2), \cdots, \mathcal{N}(e_n, \sigma_n)$, respectively. Then*

$$\mathcal{M}\left\{\sum_{i=1}^{n} \xi_i x_i \leq t\right\} = \left(1 + \exp\left(\frac{\pi\left(\sum_{i=1}^{n} e_i x_i - t\right)}{\sqrt{3}\sum_{i=1}^{n} \sigma_i x_i}\right)\right)^{-1}. \qquad (2.43)$$

Proof: Since $\xi_1, \xi_2, \cdots, \xi_n$ are independently normal uncertain variables, their weighted sum $\xi_1 x_1 + \xi_2 x_2 + \cdots + \xi_n x_n$ is also a normal uncertain variable

$$\mathcal{N}\left(\sum_{i=1}^{n} e_i x_i, \sum_{i=1}^{n} \sigma_i x_i\right).$$

From this fact we may derive the result immediately.

Theorem 2.11 *Assume $x_1, x_2, \cdots, x_n$ are nonnegative decision variables, and $\xi_1, \xi_2, \cdots, \xi_n$ are independently lognormal uncertain variables $\mathcal{LOGN}(e_1, \sigma_1)$, $\mathcal{LOGN}(e_2, \sigma_2), \cdots, \mathcal{LOGN}(e_n, \sigma_n)$, respectively. Then*

$$\mathcal{M}\left\{\sum_{i=1}^{n} \xi_i x_i \leq t\right\} = \Psi(t) \qquad (2.44)$$

where Ψ is determined by

$$\Psi^{-1}(\alpha) = \sum_{i=1}^{n} \exp(e_i)\left(\frac{\alpha}{1-\alpha}\right)^{\sqrt{3}\sigma_i/\pi} x_i. \qquad (2.45)$$

Proof: Since $\xi_1, \xi_2, \cdots, \xi_n$ are independently lognormal uncertain variables, the uncertainty distribution Ψ of $\xi_1 x_1 + \xi_2 x_2 + \cdots + \xi_n x_n$ is just determined by (2.45). From this fact we may derive the result immediately.

2.5 Uncertain Dynamic Programming

In order to model uncertain decision processes, Liu [129] proposed a general framework of uncertain dynamic programming, including expected value dynamic programming, chance-constrained dynamic programming as well as dependent-chance dynamic programming.

Expected Value Dynamic Programming

Consider an N-stage decision system in which $(\boldsymbol{a}_1, \boldsymbol{a}_2, \cdots, \boldsymbol{a}_N)$ represents the state vector, $(\boldsymbol{x}_1, \boldsymbol{x}_2, \cdots, \boldsymbol{x}_N)$ the decision vector, $(\boldsymbol{\xi}_1, \boldsymbol{\xi}_2, \cdots, \boldsymbol{\xi}_N)$ the uncertain vector. We also assume that the state transition function is

$$\boldsymbol{a}_{n+1} = T(\boldsymbol{a}_n, \boldsymbol{x}_n, \boldsymbol{\xi}_n), \quad n = 1, 2, \cdots, N-1. \qquad (2.46)$$

In order to maximize the expected return over the horizon, we may use the following expected value dynamic programming,

$$\begin{cases} f_N(\boldsymbol{a}) = \max_{E[g_N(\boldsymbol{a},\boldsymbol{x},\boldsymbol{\xi}_N)]\leq 0} E[r_N(\boldsymbol{a},\boldsymbol{x},\boldsymbol{\xi}_N)] \\ f_n(\boldsymbol{a}) = \max_{E[g_n(\boldsymbol{a},\boldsymbol{x},\boldsymbol{\xi}_n)]\leq 0} E[r_n(\boldsymbol{a},\boldsymbol{x},\boldsymbol{\xi}_n) + f_{n+1}(T(\boldsymbol{a},\boldsymbol{x},\boldsymbol{\xi}_n))] \\ n \leq N-1 \end{cases} \qquad (2.47)$$

where r_n are the return functions at the nth stages, $n = 1, 2, \cdots, N$, respectively.

Chance-Constrained Dynamic Programming

In order to maximize the optimistic return over the horizon, we may use the following chance-constrained dynamic programming,

$$\begin{cases} f_N(\boldsymbol{a}) = \max_{\mathcal{M}\{g_N(\boldsymbol{a},\boldsymbol{x},\boldsymbol{\xi}_N)\leq 0\}\geq\alpha} \overline{r}_N(\boldsymbol{a},\boldsymbol{x},\boldsymbol{\xi}_N) \\ f_n(\boldsymbol{a}) = \max_{\mathcal{M}\{g_n(\boldsymbol{a},\boldsymbol{x},\boldsymbol{\xi}_n)\leq 0\}\geq\alpha} \overline{r}_n(\boldsymbol{a},\boldsymbol{x},\boldsymbol{\xi}_n) + f_{n+1}(T(\boldsymbol{a},\boldsymbol{x},\boldsymbol{\xi}_n)) \\ n \leq N-1 \end{cases} \qquad (2.48)$$

where the functions $\overline{r}_n$ are defined by

$$\overline{r}_n(\boldsymbol{a},\boldsymbol{x},\boldsymbol{\xi}_n) = \sup\left\{\overline{r} \mid \mathcal{M}\{r_n(\boldsymbol{a},\boldsymbol{x},\boldsymbol{\xi}_n) \geq \overline{r}\} \geq \beta\right\} \qquad (2.49)$$

for $n = 1, 2, \cdots, N$. If we want to maximize the pessimistic return over the horizon, then we must define the functions $\overline{r}_n$ as

$$\overline{r}_n(\boldsymbol{a},\boldsymbol{x},\boldsymbol{\xi}_n) = \inf\left\{\overline{r} \mid \mathcal{M}\{r_n(\boldsymbol{a},\boldsymbol{x},\boldsymbol{\xi}_n) \leq \overline{r}\} \geq \beta\right\} \qquad (2.50)$$

for $n = 1, 2, \cdots, N$.

Dependent-Chance Dynamic Programming

In order to maximize the chance over the horizon, we may employ the following dependent-chance dynamic programming,

$$\begin{cases} f_N(\boldsymbol{a}) = \max_{g_N(\boldsymbol{a},\boldsymbol{x},\boldsymbol{\xi}_N)\leq 0} \mathcal{M}\{h_N(\boldsymbol{a},\boldsymbol{x},\boldsymbol{\xi}_N) \leq 0\} \\ f_n(\boldsymbol{a}) = \max_{g_n(\boldsymbol{a},\boldsymbol{x},\boldsymbol{\xi}_n)\leq 0} \mathcal{M}\{h_n(\boldsymbol{a},\boldsymbol{x},\boldsymbol{\xi}_n) \leq 0\} + f_{n+1}(T(\boldsymbol{a},\boldsymbol{x},\boldsymbol{\xi}_n)) \\ n \leq N-1 \end{cases}$$

where $h_n(\boldsymbol{a},\boldsymbol{x},\boldsymbol{\xi}_n) \leq 0$ are the events, and $g_n(\boldsymbol{a},\boldsymbol{x},\boldsymbol{\xi}_n) \leq 0$ are the uncertain environments at the nth stages, $n = 1, 2, \cdots, N$, respectively.

2.6 Uncertain Multilevel Programming

In order to model uncertain decentralized decision systems, Liu [129] presented three types of uncertain multilevel programming, including expected value multilevel programming, chance-constrained multilevel programming and dependent-chance multilevel programming, and provided the concept of Stackelberg-Nash equilibrium to uncertain multilevel programming.

Expected Value Multilevel Programming

Assume that in a decentralized two-level decision system there is one leader and m followers. Let $\boldsymbol{x}$ and $\boldsymbol{y}_i$ be the control vectors of the leader and the ith followers, $i = 1, 2, \cdots, m$, respectively. We also assume that the objective functions of the leader and ith followers are $F(\boldsymbol{x}, \boldsymbol{y}_1, \cdots, \boldsymbol{y}_m, \boldsymbol{\xi})$ and $f_i(\boldsymbol{x}, \boldsymbol{y}_1, \cdots, \boldsymbol{y}_m, \boldsymbol{\xi})$, $i = 1, 2, \cdots, m$, respectively, where $\boldsymbol{\xi}$ is an uncertain vector.

Let the feasible set of control vector $\boldsymbol{x}$ of the leader be defined by the expected constraint

$$E[G(\boldsymbol{x}, \boldsymbol{\xi})] \leq 0 \tag{2.51}$$

where G is a vector-valued function and 0 is a zero vector. Then for each decision $\boldsymbol{x}$ chosen by the leader, the feasibility of control vectors $\boldsymbol{y}_i$ of the ith followers should be dependent on not only $\boldsymbol{x}$ but also $\boldsymbol{y}_1, \cdots, \boldsymbol{y}_{i-1}, \boldsymbol{y}_{i+1}, \cdots, \boldsymbol{y}_m$, and generally represented by the expected constraints,

$$E[g_i(\boldsymbol{x}, \boldsymbol{y}_1, \boldsymbol{y}_2, \cdots, \boldsymbol{y}_m, \boldsymbol{\xi})] \leq 0 \tag{2.52}$$

where g_i are vector-valued functions, $i = 1, 2, \cdots, m$, respectively.

Assume that the leader first chooses his control vector $\boldsymbol{x}$, and the followers determine their control array $(\boldsymbol{y}_1, \boldsymbol{y}_2, \cdots, \boldsymbol{y}_m)$ after that. In order to maximize the expected objective of the leader, we have the following expected value bilevel programming,

$$
\begin{cases}
\max_{\boldsymbol{x}} E[F(\boldsymbol{x}, \boldsymbol{y}_1^*, \boldsymbol{y}_2^*, \cdots, \boldsymbol{y}_m^*, \boldsymbol{\xi})] \\
\text{subject to:} \\
\quad E[G(\boldsymbol{x}, \boldsymbol{\xi})] \leq 0 \\
\quad (\boldsymbol{y}_1^*, \boldsymbol{y}_2^*, \cdots, \boldsymbol{y}_m^*) \text{ solves problems } (i = 1, 2, \cdots, m) \\
\quad \begin{cases}
\max_{\boldsymbol{y}_i} E[f_i(\boldsymbol{x}, \boldsymbol{y}_1, \boldsymbol{y}_2, \cdots, \boldsymbol{y}_m, \boldsymbol{\xi})] \\
\text{subject to:} \\
\quad E[g_i(\boldsymbol{x}, \boldsymbol{y}_1, \boldsymbol{y}_2, \cdots, \boldsymbol{y}_m, \boldsymbol{\xi})] \leq 0.
\end{cases}
\end{cases}
\tag{2.53}
$$

Definition 2.11 *Let $\boldsymbol{x}$ be a feasible control vector of the leader. A Nash equilibrium of followers is the feasible array $(\boldsymbol{y}_1^*, \boldsymbol{y}_2^*, \cdots, \boldsymbol{y}_m^*)$ with respect to $\boldsymbol{x}$ if*

$$
\begin{aligned}
E[f_i(\boldsymbol{x}, \boldsymbol{y}_1^*, \cdots, \boldsymbol{y}_{i-1}^*, \boldsymbol{y}_i, \boldsymbol{y}_{i+1}^*, \cdots, \boldsymbol{y}_m^*, \boldsymbol{\xi})] \\
\leq E[f_i(\boldsymbol{x}, \boldsymbol{y}_1^*, \cdots, \boldsymbol{y}_{i-1}^*, \boldsymbol{y}_i^*, \boldsymbol{y}_{i+1}^*, \cdots, \boldsymbol{y}_m^*, \boldsymbol{\xi})]
\end{aligned}
\tag{2.54}
$$

for any feasible array $(\boldsymbol{y}_1^*, \cdots, \boldsymbol{y}_{i-1}^*, \boldsymbol{y}_i, \boldsymbol{y}_{i+1}^*, \cdots, \boldsymbol{y}_m^*)$ and $i = 1, 2, \cdots, m$.

Definition 2.12 *Suppose that $\boldsymbol{x}^*$ is a feasible control vector of the leader and $(\boldsymbol{y}_1^*, \boldsymbol{y}_2^*, \cdots, \boldsymbol{y}_m^*)$ is a Nash equilibrium of followers with respect to $\boldsymbol{x}^*$. We call the array $(\boldsymbol{x}^*, \boldsymbol{y}_1^*, \boldsymbol{y}_2^*, \cdots, \boldsymbol{y}_m^*)$ a Stackelberg-Nash equilibrium to the expected value bilevel programming (2.53) if*

$$E[F(\overline{\boldsymbol{x}}, \overline{\boldsymbol{y}}_1, \overline{\boldsymbol{y}}_2, \cdots, \overline{\boldsymbol{y}}_m, \boldsymbol{\xi})] \leq E[F(\boldsymbol{x}^*, \boldsymbol{y}_1^*, \boldsymbol{y}_2^*, \cdots, \boldsymbol{y}_m^*, \boldsymbol{\xi})] \qquad (2.55)$$

for any feasible control vector $\overline{\boldsymbol{x}}$ and the Nash equilibrium $(\overline{\boldsymbol{y}}_1, \overline{\boldsymbol{y}}_2, \cdots, \overline{\boldsymbol{y}}_m)$ with respect to $\overline{\boldsymbol{x}}$.

Chance-Constrained Multilevel Programming

In order to maximize the optimistic return subject to the chance constraint, we may use the following chance-constrained bilevel programming,

$$\begin{cases} \max\limits_{\boldsymbol{x}} \max\limits_{\overline{F}} \overline{F} \\ \text{subject to:} \\ \qquad \mathcal{M}\{F(\boldsymbol{x}, \boldsymbol{y}_1^*, \boldsymbol{y}_2^*, \cdots, \boldsymbol{y}_m^*, \boldsymbol{\xi}) \geq \overline{F}\} \geq \beta \\ \qquad \mathcal{M}\{G(\boldsymbol{x}, \boldsymbol{\xi}) \leq 0\} \geq \alpha \\ \qquad (\boldsymbol{y}_1^*, \boldsymbol{y}_2^*, \cdots, \boldsymbol{y}_m^*) \text{ solves problems } (i = 1, 2, \cdots, m) \\ \qquad \begin{cases} \max\limits_{\boldsymbol{y_i}} \max\limits_{\overline{f}_i} \overline{f}_i \\ \text{subject to:} \\ \qquad \mathcal{M}\{f_i(\boldsymbol{x}, \boldsymbol{y}_1, \boldsymbol{y}_2, \cdots, \boldsymbol{y}_m, \boldsymbol{\xi}) \geq \overline{f}_i\} \geq \beta_i \\ \qquad \mathcal{M}\{g_i(\boldsymbol{x}, \boldsymbol{y}_1, \boldsymbol{y}_2, \cdots, \boldsymbol{y}_m, \boldsymbol{\xi}) \leq 0\} \geq \alpha_i \end{cases} \end{cases} \qquad (2.56)$$

where $\alpha, \beta, \alpha_i, \beta_i$, $i = 1, 2, \cdots, m$ are predetermined confidence levels.

Definition 2.13 *Let $\boldsymbol{x}$ be a feasible control vector of the leader. A Nash equilibrium of followers is the feasible array $(\boldsymbol{y}_1^*, \boldsymbol{y}_2^*, \cdots, \boldsymbol{y}_m^*)$ with respect to $\boldsymbol{x}$ if*

$$\begin{aligned} \overline{f}_i(\boldsymbol{x}, \boldsymbol{y}_1^*, \cdots, \boldsymbol{y}_{i-1}^*, \boldsymbol{y}_i, \boldsymbol{y}_{i+1}^*, \cdots, \boldsymbol{y}_m^*) \\ \leq \overline{f}_i(\boldsymbol{x}, \boldsymbol{y}_1^*, \cdots, \boldsymbol{y}_{i-1}^*, \boldsymbol{y}_i^*, \boldsymbol{y}_{i+1}^*, \cdots, \boldsymbol{y}_m^*) \end{aligned} \qquad (2.57)$$

for any feasible array $(\boldsymbol{y}_1^*, \cdots, \boldsymbol{y}_{i-1}^*, \boldsymbol{y}_i, \boldsymbol{y}_{i+1}^*, \cdots, \boldsymbol{y}_m^*)$ and $i = 1, 2, \cdots, m$.

Definition 2.14 *Suppose that $\boldsymbol{x}^*$ is a feasible control vector of the leader and $(\boldsymbol{y}_1^*, \boldsymbol{y}_2^*, \cdots, \boldsymbol{y}_m^*)$ is a Nash equilibrium of followers with respect to $\boldsymbol{x}^*$. The array $(\boldsymbol{x}^*, \boldsymbol{y}_1^*, \boldsymbol{y}_2^*, \cdots, \boldsymbol{y}_m^*)$ is called a Stackelberg-Nash equilibrium to the chance-constrained bilevel programming (2.56) if*

$$\overline{F}(\overline{\boldsymbol{x}}, \overline{\boldsymbol{y}}_1, \overline{\boldsymbol{y}}_2, \cdots, \overline{\boldsymbol{y}}_m) \leq \overline{F}(\boldsymbol{x}^*, \boldsymbol{y}_1^*, \boldsymbol{y}_2^*, \cdots, \boldsymbol{y}_m^*) \qquad (2.58)$$

for any feasible control vector $\overline{\boldsymbol{x}}$ and the Nash equilibrium $(\overline{\boldsymbol{y}}_1, \overline{\boldsymbol{y}}_2, \cdots, \overline{\boldsymbol{y}}_m)$ with respect to $\overline{\boldsymbol{x}}$.

In order to maximize the pessimistic return, we have the following mini-max chance-constrained bilevel programming,

$$
\begin{cases}
\max\limits_{\boldsymbol{x}} \min\limits_{\overline{F}} \overline{F} \\
\text{subject to:} \\
\quad \mathcal{M}\{F(\boldsymbol{x}, \boldsymbol{y}_1^*, \boldsymbol{y}_2^*, \cdots, \boldsymbol{y}_m^*, \boldsymbol{\xi}) \leq \overline{F}\} \geq \beta \\
\quad \mathcal{M}\{G(\boldsymbol{x}, \boldsymbol{\xi}) \leq 0\} \geq \alpha \\
\quad (\boldsymbol{y}_1^*, \boldsymbol{y}_2^*, \cdots, \boldsymbol{y}_m^*) \text{ solves problems } (i = 1, 2, \cdots, m) \\
\quad \begin{cases}
\max\limits_{\boldsymbol{y_i}} \min\limits_{\overline{f}_i} \overline{f}_i \\
\text{subject to:} \\
\quad \mathcal{M}\{f_i(\boldsymbol{x}, \boldsymbol{y}_1, \boldsymbol{y}_2, \cdots, \boldsymbol{y}_m, \boldsymbol{\xi}) \leq \overline{f}_i\} \geq \beta_i \\
\quad \mathcal{M}\{g_i(\boldsymbol{x}, \boldsymbol{y}_1, \boldsymbol{y}_2, \cdots, \boldsymbol{y}_m, \boldsymbol{\xi}) \leq 0\} \geq \alpha_i.
\end{cases}
\end{cases} \tag{2.59}
$$

Dependent-Chance Multilevel Programming

Let $H(\boldsymbol{x}, \boldsymbol{y}_1, \boldsymbol{y}_2, \cdots, \boldsymbol{y}_m, \boldsymbol{\xi}) \leq 0$ and $h_i(\boldsymbol{x}, \boldsymbol{y}_1, \boldsymbol{y}_2, \cdots, \boldsymbol{y}_m, \boldsymbol{\xi}) \leq 0$ be the tasks of the leader and ith followers, $i = 1, 2, \cdots, m$, respectively. In order to maximize the chance functions of the leader and followers, we have the following dependent-chance bilevel programming,

$$
\begin{cases}
\max\limits_{\boldsymbol{x}} \mathcal{M}\{H(\boldsymbol{x}, \boldsymbol{y}_1^*, \boldsymbol{y}_2^*, \cdots, \boldsymbol{y}_m^*, \boldsymbol{\xi}) \leq 0\} \\
\text{subject to:} \\
\quad G(\boldsymbol{x}, \boldsymbol{\xi}) \leq 0 \\
\quad (\boldsymbol{y}_1^*, \boldsymbol{y}_2^*, \cdots, \boldsymbol{y}_m^*) \text{ solves problems } (i = 1, 2, \cdots, m) \\
\quad \begin{cases}
\max\limits_{\boldsymbol{y_i}} \mathcal{M}\{h_i(\boldsymbol{x}, \boldsymbol{y}_1, \boldsymbol{y}_2, \cdots, \boldsymbol{y}_m, \boldsymbol{\xi}) \leq 0\} \\
\text{subject to:} \\
\quad g_i(\boldsymbol{x}, \boldsymbol{y}_1, \boldsymbol{y}_2, \cdots, \boldsymbol{y}_m, \boldsymbol{\xi}) \leq 0.
\end{cases}
\end{cases} \tag{2.60}
$$

Definition 2.15 *Let $\boldsymbol{x}$ be a control vector of the leader. We call the array $(\boldsymbol{y}_1^*, \boldsymbol{y}_2^*, \cdots, \boldsymbol{y}_m^*)$ a Nash equilibrium of followers with respect to $\boldsymbol{x}$ if*

$$
\begin{aligned}
&\mathcal{M}\{h_i(\boldsymbol{x}, \boldsymbol{y}_1^*, \cdots, \boldsymbol{y}_{i-1}^*, \boldsymbol{y}_i, \boldsymbol{y}_{i+1}^*, \cdots, \boldsymbol{y}_m^*, \boldsymbol{\xi}) \leq 0\} \\
&\quad \leq \mathcal{M}\{h_i(\boldsymbol{x}, \boldsymbol{y}_1^*, \cdots, \boldsymbol{y}_{i-1}^*, \boldsymbol{y}_i^*, \boldsymbol{y}_{i+1}^*, \cdots, \boldsymbol{y}_m^*, \boldsymbol{\xi}) \leq 0\}
\end{aligned} \tag{2.61}
$$

subject to the uncertain environment $g_i(\boldsymbol{x}, \boldsymbol{y}_1, \boldsymbol{y}_2, \cdots, \boldsymbol{y}_m, \boldsymbol{\xi}) \leq 0, i = 1, 2, \cdots, m$ for any array $(\boldsymbol{y}_1^*, \cdots, \boldsymbol{y}_{i-1}^*, \boldsymbol{y}_i, \boldsymbol{y}_{i+1}^*, \cdots, \boldsymbol{y}_m^*)$ and $i = 1, 2, \cdots, m$.

Definition 2.16 *Let $\boldsymbol{x}^*$ be a control vector of the leader, and $(\boldsymbol{y}_1^*, \boldsymbol{y}_2^*, \cdots, \boldsymbol{y}_m^*)$ a Nash equilibrium of followers with respect to $\boldsymbol{x}^*$. Then $(\boldsymbol{x}^*, \boldsymbol{y}_1^*, \boldsymbol{y}_2^*, \cdots, \boldsymbol{y}_m^*)$ is called a Stackelberg-Nash equilibrium to the dependent-chance bilevel programming (2.60) if*

$$
\mathcal{M}\{H(\overline{\boldsymbol{x}}, \overline{\boldsymbol{y}}_1, \overline{\boldsymbol{y}}_2, \cdots, \overline{\boldsymbol{y}}_m, \boldsymbol{\xi}) \leq 0\} \leq \mathcal{M}\{H(\boldsymbol{x}^*, \boldsymbol{y}_1^*, \boldsymbol{y}_2^*, \cdots, \boldsymbol{y}_m^*, \boldsymbol{\xi}) \leq 0\}
$$

subject to the uncertain environment $G(\boldsymbol{x}, \boldsymbol{\xi}) \leq 0$ for any control vector $\overline{\boldsymbol{x}}$ and the Nash equilibrium $(\overline{\boldsymbol{y}}_1, \overline{\boldsymbol{y}}_2, \cdots, \overline{\boldsymbol{y}}_m)$ with respect to $\overline{\boldsymbol{x}}$.

2.7   System Reliability Design

One of the approaches to improve system reliability is to provide redundancy for components in a system. There are two ways to provide component redundancy: parallel redundancy and standby redundancy. In parallel redundancy, all redundant elements are required to operate simultaneously. This method is usually used when element replacements are not permitted during the system operation. In standby redundancy, one of the redundant elements begins to work only when the active element fails. This method is usually employed when the replacement is allowable and can be finished immediately. The system reliability design problem is to determine the optimal number of redundant elements for each component so as to optimize some system performance.

System reliability design problem has been well studied. Modelling system reliability design problem via uncertain programming was started by Liu [129] in 2009. This section provides an uncertain programming model for system reliability design problem.

Assume that a system consists of n components, and the ith components consist of x_i redundant elements, $i = 1, 2, \cdots, n$, respectively. For example, Figure 2.1 shows a bridge system in which we suppose that redundant elements are in standby for the first and second components, and are in parallel for the third to fifth components.
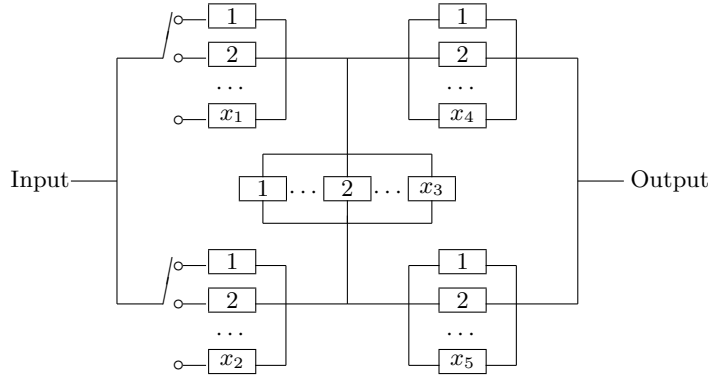
Figure 2.1: A Bridge System

The first problem is how to estimate the system lifetime when the value of the vector $\boldsymbol{x} = (x_1, x_2, \cdots, x_n)$ is determined. For such a decision vector $\boldsymbol{x}$,

suppose that the redundant elements j operating in components i have lifetimes ξ_{ij} with uncertainty distributions Φ_{ij}, $j = 1, 2, \cdots, x_i$, $i = 1, 2, \cdots, n$, respectively. Each element lifetime ξ_{ij} is represented by the 99 method as

$$
\begin{array}{|c|c|c|c|c|}
\hline
0.01 & 0.02 & 0.03 & \cdots & 0.99 \\
\hline
t_{ij}^1 & t_{ij}^2 & t_{ij}^3 & \cdots & t_{ij}^{99} \\
\hline
\end{array}
\tag{2.62}
$$

For convenience, we use the vector

$$
\boldsymbol{\xi} = (\xi_{11}, \xi_{12}, \cdots, \xi_{1x_1}, \xi_{21}, \xi_{22}, \cdots, \xi_{2x_2}, \cdots, \xi_{n1}, \xi_{n2}, \cdots, \xi_{nx_n})
\tag{2.63}
$$

to denote the lifetimes of all redundant elements in the system.

Standby Component Lifetimes

For a standby redundancy component i (1 or 2), the lifetime η_i is the sum of lifetimes of its elements,

$$
\eta_i(\boldsymbol{x}, \boldsymbol{\xi}) = \sum_{j=1}^{x_i} \xi_{ij}
\tag{2.64}
$$

whose uncertainty distribution is

$$
\Psi_i(\boldsymbol{x}; t) = \sup_{t_1 + t_2 + \cdots + t_{x_i} = t} \min_{1 \le j \le x_i} \Phi_{ij}(t_j).
\tag{2.65}
$$

Thus the 99 method yields that the standby component lifetime $\eta_i(\boldsymbol{x}, \boldsymbol{\xi})$ is

$$
\begin{array}{|c|c|c|c|c|}
\hline
0.01 & 0.02 & 0.03 & \cdots & 0.99 \\
\hline
\displaystyle\sum_{j=1}^{x_i} t_{ij}^1 & \displaystyle\sum_{j=1}^{x_i} t_{ij}^2 & \displaystyle\sum_{j=1}^{x_i} t_{ij}^3 & \cdots & \displaystyle\sum_{j=1}^{x_i} t_{ij}^{99} \\
\hline
\end{array}
\tag{2.66}
$$

Parallel Component Lifetimes

For a parallel redundancy component i (3, 4 or 5), the lifetime η_i is the maximum of lifetimes of its elements,

$$
\eta_i(\boldsymbol{x}, \boldsymbol{\xi}) = \max_{1 \le j \le x_i} \xi_{ij}
\tag{2.67}
$$

whose uncertainty distribution is

$$
\Psi_i(\boldsymbol{x}; t) = \min_{1 \le j \le x_i} \Phi_{ij}(t).
\tag{2.68}
$$

Thus the 99 method yields that the parallel component lifetime $\eta_i(\boldsymbol{x}, \boldsymbol{\xi})$ is

$$
\begin{array}{|c|c|c|c|c|}
\hline
0.01 & 0.02 & 0.03 & \cdots & 0.99 \\
\hline
\displaystyle\bigvee_{j=1}^{x_i} t_{ij}^1 & \displaystyle\bigvee_{j=1}^{x_i} t_{ij}^2 & \displaystyle\bigvee_{j=1}^{x_i} t_{ij}^3 & \cdots & \displaystyle\bigvee_{j=1}^{x_i} t_{ij}^{99} \\
\hline
\end{array}
\tag{2.69}
$$

Path Lifetimes

For the bridge system shown in Figure 2.1, there are 4 paths from the input of the system to the output:

$$\text{Path 1: input-1-4-output,}$$
$$\text{Path 2: input-2-5-output,}$$
$$\text{Path 3: input-2-3-4-output,}$$
$$\text{Path 4: input-1-3-5-output.}$$

Assume each path works if and only if all components on which are working. Then the lifetime of each path is the minimum of lifetimes of its components. It follows that the lifetime of Path 1 (input-1-4-output) is

$$\tau_1(\boldsymbol{x}, \boldsymbol{\xi}) = \eta_1(\boldsymbol{x}, \boldsymbol{\xi}) \wedge \eta_4(\boldsymbol{x}, \boldsymbol{\xi}) \tag{2.70}$$

whose uncertainty distribution is

$$\Upsilon_1(\boldsymbol{x}; t) = \Psi_1(\boldsymbol{x}; t) \vee \Psi_4(\boldsymbol{x}; t). \tag{2.71}$$

Thus the 99 method yields that the lifetime $\tau_1(\boldsymbol{x}, \boldsymbol{\xi})$ is

$$\begin{array}{|c|c|c|c|}
\hline
0.01 & 0.02 & \cdots & 0.99 \\
\hline
\displaystyle\sum_{j=1}^{x_1} t_{1j}^1 \wedge \bigvee_{j=1}^{x_4} t_{4j}^1 & \displaystyle\sum_{j=1}^{x_1} t_{1j}^2 \wedge \bigvee_{j=1}^{x_4} t_{4j}^2 & \cdots & \displaystyle\sum_{j=1}^{x_1} t_{1j}^{99} \wedge \bigvee_{j=1}^{x_4} t_{4j}^{99} \\
\hline
\end{array} \tag{2.72}$$

The lifetime of Path 2 (input-2-5-output) is

$$\tau_2(\boldsymbol{x}, \boldsymbol{\xi}) = \eta_2(\boldsymbol{x}, \boldsymbol{\xi}) \wedge \eta_5(\boldsymbol{x}, \boldsymbol{\xi}) \tag{2.73}$$

whose uncertainty distribution is

$$\Upsilon_2(\boldsymbol{x}; t) = \Psi_2(\boldsymbol{x}; t) \vee \Psi_5(\boldsymbol{x}; t). \tag{2.74}$$

Thus the 99 method yields that the lifetime $\tau_2(\boldsymbol{x}, \boldsymbol{\xi})$ is

$$\begin{array}{|c|c|c|c|}
\hline
0.01 & 0.02 & \cdots & 0.99 \\
\hline
\displaystyle\sum_{j=1}^{x_2} t_{2j}^1 \wedge \bigvee_{j=1}^{x_5} t_{5j}^1 & \displaystyle\sum_{j=1}^{x_2} t_{2j}^2 \wedge \bigvee_{j=1}^{x_5} t_{5j}^2 & \cdots & \displaystyle\sum_{j=1}^{x_2} t_{2j}^{99} \wedge \bigvee_{j=1}^{x_5} t_{5j}^{99} \\
\hline
\end{array} \tag{2.75}$$

The lifetime of Path 3 (input-2-3-4-output) is

$$\tau_3(\boldsymbol{x}, \boldsymbol{\xi}) = \eta_2(\boldsymbol{x}, \boldsymbol{\xi}) \wedge \eta_3(\boldsymbol{x}, \boldsymbol{\xi}) \wedge \eta_4(\boldsymbol{x}, \boldsymbol{\xi}) \tag{2.76}$$

whose uncertainty distribution is

$$\Upsilon_3(\boldsymbol{x}; t) = \Psi_2(\boldsymbol{x}; t) \vee \Psi_3(\boldsymbol{x}; t) \vee \Psi_4(\boldsymbol{x}, t). \tag{2.77}$$

Thus the 99 method yields that the lifetime $\tau_3(\boldsymbol{x}, \boldsymbol{\xi})$ is

$$
\begin{array}{|ccc|c|ccc|}
\hline
\multicolumn{3}{|c|}{0.01} & \cdots & \multicolumn{3}{c|}{0.99} \\
\hline
\displaystyle\sum_{j=1}^{x_2} t_{2j}^1 \wedge \bigvee_{j=1}^{x_3} t_{3j}^1 \wedge \bigvee_{j=1}^{x_4} t_{4j}^1 & & & \cdots & \displaystyle\sum_{j=1}^{x_2} t_{2j}^{99} \wedge \bigvee_{j=1}^{x_3} t_{3j}^{99} \wedge \bigvee_{j=1}^{x_4} t_{4j}^{99} & & \\
\hline
\end{array}
\tag{2.78}
$$

The lifetime of Path 4 (input-1-3-5-output) is

$$
\tau_4(\boldsymbol{x}, \boldsymbol{\xi}) = \eta_1(\boldsymbol{x}, \boldsymbol{\xi}) \wedge \eta_3(\boldsymbol{x}, \boldsymbol{\xi}) \wedge \eta_5(\boldsymbol{x}, \boldsymbol{\xi})
\tag{2.79}
$$

whose uncertainty distribution is

$$
\Upsilon_4(\boldsymbol{x}; t) = \Psi_1(\boldsymbol{x}; t) \vee \Psi_3(\boldsymbol{x}; t) \vee \Psi_5(\boldsymbol{x}, t)
\tag{2.80}
$$

Thus the 99 method yields that the lifetime $\tau_4(\boldsymbol{x}, \boldsymbol{\xi})$ is

$$
\begin{array}{|ccc|c|ccc|}
\hline
\multicolumn{3}{|c|}{0.01} & \cdots & \multicolumn{3}{c|}{0.99} \\
\hline
\displaystyle\sum_{j=1}^{x_1} t_{1j}^1 \wedge \bigvee_{j=1}^{x_3} t_{3j}^1 \wedge \bigvee_{j=1}^{x_5} t_{5j}^1 & & & \cdots & \displaystyle\sum_{j=1}^{x_1} t_{1j}^{99} \wedge \bigvee_{j=1}^{x_3} t_{3j}^{99} \wedge \bigvee_{j=1}^{x_5} t_{5j}^{99} & & \\
\hline
\end{array}
\tag{2.81}
$$

System Lifetime

How do we calculate the system lifetime $T(\boldsymbol{x}, \boldsymbol{\xi})$? It is problem-dependent. Assume the system works if and only if there is a path of working components. Then the system lifetime is

$$
T(\boldsymbol{x}, \boldsymbol{\xi}) = \tau_1(\boldsymbol{x}, \boldsymbol{\xi}) \vee \tau_2(\boldsymbol{x}, \boldsymbol{\xi}) \vee \tau_3(\boldsymbol{x}, \boldsymbol{\xi}) \vee \tau_4(\boldsymbol{x}, \boldsymbol{\xi})
\tag{2.82}
$$

whose uncertainty distribution is

$$
\Upsilon(\boldsymbol{x}; t) = \Upsilon_1(\boldsymbol{x}; t) \wedge \Upsilon_2(\boldsymbol{x}; t) \wedge \Upsilon_3(\boldsymbol{x}; t) \wedge \Upsilon_4(\boldsymbol{x}; t).
\tag{2.83}
$$

The 99 method yields that the system lifetime $T(\boldsymbol{x}, \boldsymbol{\xi})$ is

$$
\begin{array}{|c|c|c|c|c|}
\hline
0.01 & 0.02 & 0.03 & \cdots & 0.99 \\
\hline
s_1 & s_2 & s_3 & \cdots & s_{99} \\
\hline
\end{array}
\tag{2.84}
$$

where

$$
s_k = \max \left\{
\begin{array}{c}
\displaystyle\sum_{j=1}^{x_1} t_{1j}^k \wedge \bigvee_{j=1}^{x_4} t_{4j}^k \\
\displaystyle\sum_{j=1}^{x_2} t_{2j}^k \wedge \bigvee_{j=1}^{x_5} t_{5j}^k \\
\displaystyle\sum_{j=1}^{x_2} t_{2j}^k \wedge \bigvee_{j=1}^{x_3} t_{3j}^k \wedge \bigvee_{j=1}^{x_4} t_{4j}^k \\
\displaystyle\sum_{j=1}^{x_1} t_{1j}^k \wedge \bigvee_{j=1}^{x_3} t_{3j}^k \wedge \bigvee_{j=1}^{x_5} t_{5j}^k
\end{array}
\right\}
\tag{2.85}
$$

for $k = 1, 2, \cdots, 99$.

System Reliability Model

By *system reliability* we mean $\mathcal{M}\{T(\boldsymbol{x}, \boldsymbol{\xi}) \geq T^0\}$, i.e., the uncertain measure that the system lifetime is greater than or equal to the given time T^0. If we want to maximize the system reliability under a cost constraint, then we use the following system reliability model,

$$
\begin{cases}
\max \mathcal{M}\{T(\boldsymbol{x}, \boldsymbol{\xi}) \geq T^0\} \\
\text{subject to:} \\
\quad C(\boldsymbol{x}) \leq \overline{C} \\
\quad \boldsymbol{x} \geq 1, \text{ integer vector.}
\end{cases} \tag{2.86}
$$

When $\Upsilon(\boldsymbol{x}; t)$ determined by (2.83) is a continuous function with respect to t, the above model is equivalent to

$$
\begin{cases}
\max 1 - \Upsilon(\boldsymbol{x}; T^0) \\
\text{subject to:} \\
\quad C(\boldsymbol{x}) \leq \overline{C} \\
\quad \boldsymbol{x} \geq 1, \text{ integer vector}
\end{cases} \tag{2.87}
$$

where $\Upsilon(\boldsymbol{x}; T^0)$ may be approximately estimated by

$$
\Upsilon(\boldsymbol{x}; T^0) = \frac{k}{100}, \quad \text{if } s_k \leq T^0 < s_{k+1} \text{ for some } k \tag{2.88}
$$

and $s_1, s_2, \cdots, s_{99}$ are given by the 99 method in (2.84).

Numerical Experiment

Assume that all the element lifetimes are lognormal uncertain variables,

$$
\xi_{ij} \sim \mathcal{LOGN}(i, 1), \quad j = 1, 2, \cdots, x_i, \quad i = 1, 2, \cdots, 5.
$$

We also assume that the costs of one element in components i are

$$
c_i = 100 \times i, \quad i = 1, 2, \cdots, 5.
$$

If the total capital available is $\overline{C} = 2000$, then we have a cost constraint

$$
100x_1 + 200x_2 + 300x_3 + 400x_4 + 500x_5 \leq 2000.
$$

A run of computer program (http://orsc.edu.cn/liu/resources.htm) shows that the optimal redundancy number is

$$
\boldsymbol{x}^* = (1, 3, 1, 1, 1)
$$

whose total cost is 1900 and system reliability $\mathcal{M}\{T(\boldsymbol{x}^*, \boldsymbol{\xi}) \geq 6\}$ is 0.92 in uncertain measure.

2.8   Project Scheduling Problem

Project scheduling problem is to determine the schedule of allocating resources so as to balance the total cost and the completion time. Project scheduling problem is a hot topic. The study of project scheduling problem with uncertain factors was started by Liu [129] in 2009. This section presents an uncertain programming model for project scheduling problem.

Project scheduling is usually represented by a directed acyclic graph where nodes correspond to milestones, and arcs to activities which are basically characterized by the times and costs consumed.
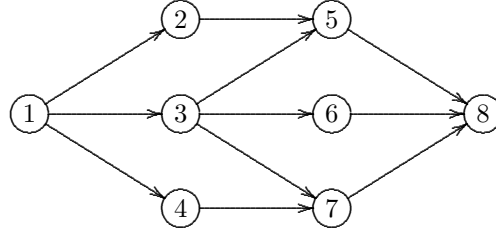
Figure 2.2: A Project Graph

Let $(\mathcal{V}, \mathcal{A})$ be a directed acyclic graph, where $\mathcal{V} = \{1, 2, \cdots, n, n+1\}$ is the set of nodes, $\mathcal{A}$ is the set of arcs, $(i,j) \in \mathcal{A}$ is the arc of the graph $(\mathcal{V}, \mathcal{A})$ from nodes i to j. It is well-known that we can rearrange the indexes of the nodes in $\mathcal{V}$ such that $i < j$ for all $(i,j) \in \mathcal{A}$.

Before we begin to study project scheduling problem with uncertain activity duration times, we first make some assumptions: (a) all of the costs needed are obtained via loans with some given interest rate; and (b) each activity can be processed only if the loan needed is allocated and all the foregoing activities are finished.

In order to model the project scheduling problem, we introduce the following indices and parameters:

ξ_{ij}: uncertain duration time of activity (i,j) in $\mathcal{A}$;

Φ_{ij}: uncertainty distribution of ξ_{ij};

c_{ij}: cost of activity (i,j) in $\mathcal{A}$;

r: interest rate;

x_i: integer decision variable representing the allocating time of all loans needed for all activities (i,j) in $\mathcal{A}$.

Starting Times

For simplicity, we write $\boldsymbol{\xi} = \{\xi_{ij} : (i,j) \in \mathcal{A}\}$, $\boldsymbol{x} = (x_1, x_2, \cdots, x_n)$, and each uncertain duration time ξ_{ij} is represented by the 99 method as

0.01	0.02	0.03	$\cdots$	0.99
t_{ij}^1	t_{ij}^2	t_{ij}^3	$\cdots$	t_{ij}^{99}

(2.89)

Let $T_i(\boldsymbol{x}, \boldsymbol{\xi})$ denote the starting time of all activities (i, j) in $\mathcal{A}$. According to the assumptions, the starting time of the total project (i.e., the starting time of of all activities $(1, j)$ in $\mathcal{A}$) should be

$$T_1(\boldsymbol{x}, \boldsymbol{\xi}) = x_1 \tag{2.90}$$

whose uncertainty distribution is

$$\Psi_1(\boldsymbol{x}; z) = \begin{cases} 0, & \text{if } z < x_1 \\ 1, & \text{if } z \geq x_1 \end{cases} \tag{2.91}$$

and is produced by the 99 method as

0.01	0.02	0.03	$\cdots$	0.99
x_1	x_1	x_1	$\cdots$	x_1

$$(2.92)$$

Generally, suppose that the starting time $T_k(\boldsymbol{x}, \boldsymbol{\xi})$ of all activities (k, j) in $\mathcal{A}$ has an uncertainty distribution $\Psi_k(\boldsymbol{x}; z)$ and is produced by the 99 method as

0.01	0.02	0.03	$\cdots$	0.99
y_k^1	y_k^2	y_k^3	$\cdots$	y_k^{99}

$$(2.93)$$

Then the starting time $T_i(\boldsymbol{x}, \boldsymbol{\xi})$ of all activities (i, j) in $\mathcal{A}$ should be

$$T_i(\boldsymbol{x}, \boldsymbol{\xi}) = x_i \vee \max_{(k,i) \in \mathcal{A}} (T_k(\boldsymbol{x}, \boldsymbol{\xi}) + \xi_{ki}) \tag{2.94}$$

whose uncertainty distribution is

$$\Psi_i(\boldsymbol{x}; z) = \begin{cases} 0, & \text{if } z < x_i \\ \min_{(k,i) \in \mathcal{A}} \sup_{z_1 + z_2 = z} \Psi_k(\boldsymbol{x}; z_1) \wedge \Phi_{ki}(z_2), & \text{if } z \geq x_i \end{cases} \tag{2.95}$$

and is produced by the 99 method as

0.01	$\cdots$	0.99
$x_i \vee \max\limits_{(k,i) \in \mathcal{A}} (y_k^1 + t_{ki}^1)$	$\cdots$	$x_i \vee \max\limits_{(k,i) \in \mathcal{A}} (y_k^{99} + t_{ki}^{99})$

$$(2.96)$$

where $y_k^1, y_k^2, \cdots, y_k^{99}$ are determined by (2.93). This recursive process may produce all starting times of activities.

Completion Time

The completion time $T(\boldsymbol{x}, \boldsymbol{\xi})$ of the total project (i.e, the finish time of all activities $(k, n+1)$ in $\mathcal{A}$) is

$$T(\boldsymbol{x}, \boldsymbol{\xi}) = \max_{(k,n+1) \in \mathcal{A}} \{T_k(\boldsymbol{x}, \boldsymbol{\xi}) + \xi_{k,n+1}\} \tag{2.97}$$

whose uncertainty distribution is

$$\Psi(\boldsymbol{x}; z) = \min_{(k,n+1)\in\mathcal{A}} \sup_{z_1+z_2=z} \Psi_k(\boldsymbol{x}; z_1) \wedge \Phi_{k,n+1}(z_2) \qquad (2.98)$$

and is produced by the 99 method as

0.01	$\cdots$	0.99
$\displaystyle\max_{(k,n+1)\in\mathcal{A}} (y_k^1 + t_{k,n+1}^1)$	$\cdots$	$\displaystyle\max_{(k,n+1)\in\mathcal{A}} (y_k^{99} + t_{k,n+1}^{99})$

$$(2.99)$$

where $y_k^1, y_k^2, \cdots, y_k^{99}$ are determined by (2.93).

Total Cost

Based on the completion time $T(\boldsymbol{x}, \boldsymbol{\xi})$, the total cost of the project can be written as

$$C(\boldsymbol{x}, \boldsymbol{\xi}) = \sum_{(i,j)\in\mathcal{A}} c_{ij} (1+r)^{\lceil T(\boldsymbol{x},\boldsymbol{\xi})-x_i \rceil} \qquad (2.100)$$

where $\lceil a \rceil$ represents the minimal integer greater than or equal to a. Note that $C(\boldsymbol{x}, \boldsymbol{\xi})$ is a discrete uncertain variable whose uncertainty distribution is described by

$$\Upsilon^{-1}(\boldsymbol{x}; \alpha) = \sum_{(i,j)\in\mathcal{A}} c_{ij} (1+r)^{\lceil \Psi^{-1}(\boldsymbol{x};\alpha)-x_i \rceil} \qquad (2.101)$$

for $0 < \alpha < 1$. Note that $T(\boldsymbol{x}, \boldsymbol{\xi})$ is obtained by the recursive process and produced by the 99 method as

0.01	0.02	0.03	$\cdots$	0.99
s_1	s_2	s_3	$\cdots$	s_{99}

$$(2.102)$$

Thus the total cost $C(\boldsymbol{x}, \boldsymbol{\xi})$ is produced by the 99 method as

0.01	$\cdots$	0.99
$\displaystyle\sum_{(i,j)\in\mathcal{A}} c_{ij} (1+r)^{\lceil s_1-x_i \rceil}$	$\cdots$	$\displaystyle\sum_{(i,j)\in\mathcal{A}} c_{ij} (1+r)^{\lceil s_{99}-x_i \rceil}$

$$(2.103)$$

Project Scheduling Model

If we want to minimize the expected cost of the project under the completion time constraint, we may construct the following project scheduling model,

$$\begin{cases} \min E[C(\boldsymbol{x}, \boldsymbol{\xi})] \\ \text{subject to:} \\ \qquad \mathcal{M}\{T(\boldsymbol{x}, \boldsymbol{\xi}) \le T^0\} \ge \alpha \\ \qquad \boldsymbol{x} \ge 0, \ \text{integer vector} \end{cases} \qquad (2.104)$$

where T^0 is a due date of the project, α is a predetermined confidence level, $T(\boldsymbol{x}, \boldsymbol{\xi})$ is the completion time defined by (2.97), and $C(\boldsymbol{x}, \boldsymbol{\xi})$ is the total cost defined by (2.100). This model is equivalent to

$$
\begin{cases}
\min \displaystyle\int_0^{+\infty} (1 - \Upsilon(\boldsymbol{x}; z))\mathrm{d}z \\
\text{subject to:} \\
\qquad \Psi(\boldsymbol{x}; T^0) \geq \alpha \\
\qquad \boldsymbol{x} \geq 0, \text{ integer vector}
\end{cases}
\tag{2.105}
$$

where Ψ is determined by (2.98) and Υ is determined by (2.101). Note that the completion time $T(\boldsymbol{x}, \boldsymbol{\xi})$ and total cost $C(\boldsymbol{x}, \boldsymbol{\xi})$ are obtained by the recursive process and are respectively represented by the 99 method as follows,

0.01	0.02	0.03	$\cdots$	0.99
s_1	s_2	s_3	$\cdots$	s_{99}

0.01	0.02	0.03	$\cdots$	0.99
c_1	c_2	c_3	$\cdots$	c_{99}

$$\tag{2.106}$$

Thus the project scheduling model is simplified as follows,

$$
\begin{cases}
\min \ (c_1 + c_2 + \cdots + c_{99})/99 \\
\text{subject to:} \\
\qquad k/100 \geq \alpha \text{ if } s_k \geq T^0 \\
\qquad \boldsymbol{x} \geq 0, \text{ integer vector.}
\end{cases}
\tag{2.107}
$$

Numerical Experiment

Assume that all duration times of activities are linear uncertain variables,

$$\xi_{ij} \sim \mathcal{L}(3i, 3j), \quad \forall (i, j) \in \mathcal{A}$$

and assume that the costs of activities are

$$c_{ij} = i + j, \quad \forall (i, j) \in \mathcal{A}.$$

In addition, we also suppose that the interest rate is $r = 0.02$, the due date is $T^0 = 60$, and the confidence level is $\alpha = 0.85$. A run of computer program (http://orsc.edu.cn/liu/resources.htm) shows that the optimal allocating times of all loans needed for all activities are

Date	7	11	13	23	26	29
Node	1	4	3	2, 7	6	5
Loan	12	11	27	22	14	13

whose expected total cost is 166.8, and $\mathcal{M}\{T(\boldsymbol{x}^*, \boldsymbol{\xi}) \leq 60\} = 0.89$.

2.9   Vehicle Routing Problem

Vehicle routing problem (VRP) is concerned with finding efficient routes, beginning and ending at a central depot, for a fleet of vehicles to serve a number of customers.
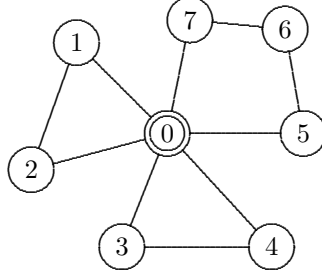
Figure 2.3: A Vehicle Routing Graph

Due to its wide applicability and economic importance, vehicle routing problem has been extensively studied. Liu [129] first introduced uncertainty theory into the research area of vehicle routing problem in 2009. In this section, vehicle routing problem will be modelled by uncertain programming.

We assume that (a) a vehicle will be assigned for only one route on which there may be more than one customer; (b) a customer will be visited by one and only one vehicle; (c) each route begins and ends at the depot; and (d) each customer specifies its time window within which the delivery is permitted or preferred to start.

Let us first introduce the following indices and model parameters:

$i = 0$: depot;

$i = 1, 2, \cdots, n$: customers;

$k = 1, 2, \cdots, m$: vehicles;

D_{ij}: travel distance from customers i to j, $i, j = 0, 1, 2, \cdots, n$;

T_{ij}: uncertain travel time from customers i to j, $i, j = 0, 1, 2, \cdots, n$;

Φ_{ij}: uncertainty distribution of T_{ij}, $i, j = 0, 1, 2, \cdots, n$;

$[a_i, b_i]$: time window of customer i, $i = 1, 2, \cdots, n$.

Operational Plan

In this book, the operational plan is represented by Liu's formulation [119] via three decision vectors $\boldsymbol{x}$, $\boldsymbol{y}$ and $\boldsymbol{t}$, where

$\boldsymbol{x} = (x_1, x_2, \cdots, x_n)$: integer decision vector representing n customers with $1 \leq x_i \leq n$ and $x_i \neq x_j$ for all $i \neq j$, $i, j = 1, 2, \cdots, n$. That is, the sequence $\{x_1, x_2, \cdots, x_n\}$ is a rearrangement of $\{1, 2, \cdots, n\}$;

$\boldsymbol{y} = (y_1, y_2, \cdots, y_{m-1})$: integer decision vector with $y_0 \equiv 0 \leq y_1 \leq y_2 \leq \cdots \leq y_{m-1} \leq n \equiv y_m$;

$\boldsymbol{t} = (t_1, t_2, \cdots, t_m)$: each t_k represents the starting time of vehicle k at the depot, $k = 1, 2, \cdots, m$.

We note that the operational plan is fully determined by the decision vectors $\boldsymbol{x}$, $\boldsymbol{y}$ and $\boldsymbol{t}$ in the following way. For each k ($1 \leq k \leq m$), if $y_k = y_{k-1}$, then vehicle k is not used; if $y_k > y_{k-1}$, then vehicle k is used and starts from the depot at time t_k, and the tour of vehicle k is $0 \rightarrow x_{y_{k-1}+1} \rightarrow x_{y_{k-1}+2} \rightarrow \cdots \rightarrow x_{y_k} \rightarrow 0$. Thus the tours of all vehicles are as follows:

Vehicle 1: $0 \rightarrow x_{y_0+1} \rightarrow x_{y_0+2} \rightarrow \cdots \rightarrow x_{y_1} \rightarrow 0$;

Vehicle 2: $0 \rightarrow x_{y_1+1} \rightarrow x_{y_1+2} \rightarrow \cdots \rightarrow x_{y_2} \rightarrow 0$;

$\cdots$

Vehicle m: $0 \rightarrow x_{y_{m-1}+1} \rightarrow x_{y_{m-1}+2} \rightarrow \cdots \rightarrow x_{y_m} \rightarrow 0$.
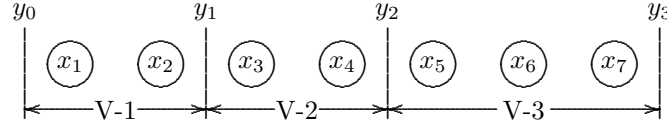
Figure 2.4: Formulation of Operational Plan

It is clear that this type of representation is intuitive, and the total number of decision variables is $n + 2m - 1$. We also note that the above decision variables $\boldsymbol{x}$, $\boldsymbol{y}$ and $\boldsymbol{t}$ ensure that: (a) each vehicle will be used at most one time; (b) all tours begin and end at the depot; (c) each customer will be visited by one and only one vehicle; and (d) there is no subtour.

Arrival Times

Let $f_i(\boldsymbol{x}, \boldsymbol{y}, \boldsymbol{t})$ be the arrival time function of some vehicles at customers i for $i = 1, 2, \cdots, n$. We remind readers that $f_i(\boldsymbol{x}, \boldsymbol{y}, \boldsymbol{t})$ are determined by the decision variables $\boldsymbol{x}$, $\boldsymbol{y}$ and $\boldsymbol{t}$, $i = 1, 2, \cdots, n$. Since unloading can start either immediately, or later, when a vehicle arrives at a customer, the calculation of $f_i(\boldsymbol{x}, \boldsymbol{y}, \boldsymbol{t})$ is heavily dependent on the operational strategy. Here we assume that the customer does not permit a delivery earlier than the time window. That is, the vehicle will wait to unload until the beginning of the time window if it arrives before the time window. If a vehicle arrives at a customer after the beginning of the time window, unloading will start immediately. For each k with $1 \leq k \leq m$, if vehicle k is used (i.e., $y_k > y_{k-1}$), then we have

$$f_{x_{y_{k-1}+1}}(\boldsymbol{x}, \boldsymbol{y}, \boldsymbol{t}) = t_k + T_{0x_{y_{k-1}+1}} \tag{2.108}$$

and

$$f_{x_{y_{k-1}+j}}(\boldsymbol{x}, \boldsymbol{y}, \boldsymbol{t}) = f_{x_{y_{k-1}+j-1}}(\boldsymbol{x}, \boldsymbol{y}, \boldsymbol{t}) \vee a_{x_{y_{k-1}+j-1}} + T_{x_{y_{k-1}+j-1}x_{y_{k-1}+j}} \tag{2.109}$$

for $2 \leq j \leq y_k - y_{k-1}$. It follows from the uncertainty of travel times T_{ij}'s that the arrival times $f_i(\boldsymbol{x}, \boldsymbol{y}, \boldsymbol{t})$, $i = 1, 2, \cdots, n$ are uncertain variables fully determined by (2.108) and (2.109).

Assume that each travel time T_{ij} from customers i to j is represented by the 99 method as

$$
\begin{array}{|c|c|c|c|c|}
\hline
0.01 & 0.02 & 0.03 & \cdots & 0.99 \\
\hline
t_{ij}^1 & t_{ij}^2 & t_{ij}^3 & \cdots & t_{ij}^{99} \\
\hline
\end{array}
\tag{2.110}
$$

If the vehicle k is used, i.e., $y_k > y_{k-1}$, then the arrival time $f_{x_{y_{k-1}+1}}(\boldsymbol{x}, \boldsymbol{y}, \boldsymbol{t})$ at the customer $x_{y_{k-1}+1}$ is an uncertain variable whose uncertainty distribution is

$$
\Psi_{x_{y_{k-1}+1}}(\boldsymbol{x}, \boldsymbol{y}, \boldsymbol{t}; z) = \Phi_{0 x_{y_{k-1}+1}}(z - t_k)
\tag{2.111}
$$

and is produced by the 99 method as

$$
\begin{array}{|c|c|c|c|}
\hline
0.01 & 0.02 & \cdots & 0.99 \\
\hline
t_k + t_{0 x_{y_{k-1}+1}}^1 & t_k + t_{0 x_{y_{k-1}+1}}^2 & \cdots & t_k + t_{0 x_{y_{k-1}+1}}^{99} \\
\hline
\end{array}
\tag{2.112}
$$

Generally, suppose that the arrival time $f_{x_{y_{k-1}+j-1}}(\boldsymbol{x}, \boldsymbol{y}, \boldsymbol{t})$ has an uncertainty distribution $\Psi_{x_{y_{k-1}+j-1}}(z)$, and is produced by the 99 method as

$$
\begin{array}{|c|c|c|c|}
\hline
0.01 & 0.02 & \cdots & 0.99 \\
\hline
s_{x_{y_{k-1}+j-1}}^1 & s_{x_{y_{k-1}+j-1}}^2 & \cdots & s_{x_{y_{k-1}+j-1}}^{99} \\
\hline
\end{array}
\tag{2.113}
$$

Note that the arrival time $f_{x_{y_{k-1}+j}}(\boldsymbol{x}, \boldsymbol{y}, \boldsymbol{t})$ at the customer $x_{y_{k-1}+j}$ has an uncertainty distribution

$$
\Psi_{x_{y_{k-1}+j}}(\boldsymbol{x}, \boldsymbol{y}, \boldsymbol{t}; z) = \sup_{z_1 + z_2 = z} \Psi^*_{x_{y_{k-1}+j-1}}(z_1) \wedge \Phi_{x_{y_{k-1}+j-1} x_{y_{k-1}+j}}(z_2)
\tag{2.114}
$$

where

$$
\Psi^*_{x_{y_{k-1}+j-1}}(z) = \begin{cases} 0, & \text{if } z < a_{x_{y_{k-1}+j-1}} \\ \Psi_{x_{y_{k-1}+j-1}}(z), & \text{if } z \geq a_{x_{y_{k-1}+j-1}} \end{cases}
\tag{2.115}
$$

for $2 \leq j \leq y_k - y_{k-1}$. Thus the arrival time $f_{x_{y_{k-1}+j}}(\boldsymbol{x}, \boldsymbol{y}, \boldsymbol{t})$ is produced by the 99 method as

$$
\begin{array}{|c|c|c|}
\hline
0.01 & \cdots & 0.99 \\
\hline
\begin{array}{c} s_{x_{y_{k-1}+j-1}}^1 \vee a_{x_{y_{k-1}+j-1}} \\ + t_{x_{y_{k-1}+j-1} x_{y_{k-1}+j}}^1 \end{array} & \cdots & \begin{array}{c} s_{x_{y_{k-1}+j-1}}^{99} \vee a_{x_{y_{k-1}+j-1}} \\ + t_{x_{y_{k-1}+j-1} x_{y_{k-1}+j}}^{99} \,1 \end{array} \\
\hline
\end{array}
\tag{2.116}
$$

where $s_{x_{y_{k-1}+j-1}}^1, s_{x_{y_{k-1}+j-1}}^2, \cdots, s_{x_{y_{k-1}+j-1}}^{99}$ are determined by (2.113). This recursive process may produce all arrival times at customers.

Travel Distance

Let $g(\boldsymbol{x}, \boldsymbol{y})$ be the total travel distance of all vehicles. Then we have

$$g(\boldsymbol{x}, \boldsymbol{y}) = \sum_{k=1}^{m} g_k(\boldsymbol{x}, \boldsymbol{y}) \qquad (2.117)$$

where

$$g_k(\boldsymbol{x}, \boldsymbol{y}) = \begin{cases} D_{0x_{y_{k-1}+1}} + \sum_{j=y_{k-1}+1}^{y_k-1} D_{x_j x_{j+1}} + D_{x_{y_k} 0}, & \text{if } y_k > y_{k-1} \\ 0, & \text{if } y_k = y_{k-1} \end{cases}$$

for $k = 1, 2, \cdots, m$.

Vehicle Routing Model

If we hope that each customer i $(1 \le i \le n)$ is visited within its time window $[a_i, b_i]$ with confidence level α_i (i.e., the vehicle arrives at customer i before time b_i), then we have the following chance constraint,

$$\mathcal{M}\{f_i(\boldsymbol{x}, \boldsymbol{y}, \boldsymbol{t}) \le b_i\} \ge \alpha_i. \qquad (2.118)$$

If we want to minimize the total travel distance of all vehicles subject to the time window constraint, then we have the following vehicle routing model,

$$\begin{cases} \min g(\boldsymbol{x}, \boldsymbol{y}) \\ \text{subject to:} \\ \quad \mathcal{M}\{f_i(\boldsymbol{x}, \boldsymbol{y}, \boldsymbol{t}) \le b_i\} \ge \alpha_i, \quad i = 1, 2, \cdots, n \\ \quad 1 \le x_i \le n, \quad i = 1, 2, \cdots, n \\ \quad x_i \ne x_j, \quad i \ne j, \ i, j = 1, 2, \cdots, n \\ \quad 0 \le y_1 \le y_2 \le \cdots \le y_{m-1} \le n \\ \quad x_i, y_j, \quad i = 1, 2, \cdots, n, \quad j = 1, 2, \cdots, m-1, \quad \text{integers} \end{cases} \qquad (2.119)$$

which is equivalent to

$$\begin{cases} \min g(\boldsymbol{x}, \boldsymbol{y}) \\ \text{subject to:} \\ \quad \Psi_i(\boldsymbol{x}, \boldsymbol{y}, \boldsymbol{t}; b_i) \ge \alpha_i, \quad i = 1, 2, \cdots, n \\ \quad 1 \le x_i \le n, \quad i = 1, 2, \cdots, n \\ \quad x_i \ne x_j, \quad i \ne j, \ i, j = 1, 2, \cdots, n \\ \quad 0 \le y_1 \le y_2 \le \cdots \le y_{m-1} \le n \\ \quad x_i, y_j, \quad i = 1, 2, \cdots, n, \quad j = 1, 2, \cdots, m-1, \quad \text{integers} \end{cases} \qquad (2.120)$$

where Ψ_i are uncertainty distributions determined by (2.111) and (2.114) for $i = 1, 2, \cdots, n$. Note that all arrival times $f_i(\boldsymbol{x}, \boldsymbol{y}, \boldsymbol{t})$, $i = 1, 2, \cdots, n$ are obtained by the 99 method and are respectively represented by

0.01	0.02	0.03	$\cdots$	0.99
s_1^1	s_1^2	s_1^3	$\cdots$	s_1^{99}

0.01	0.02	0.03	$\cdots$	0.99
s_2^1	s_2^2	s_2^3	$\cdots$	s_2^{99}

$$\vdots$$

0.01	0.02	0.03	$\cdots$	0.99
s_n^1	s_n^2	s_n^3	$\cdots$	s_n^{99}

(2.121)

Thus the vehicle routing model is simplified as follows,

$$
\begin{cases}
\min g(\boldsymbol{x}, \boldsymbol{y}) \\
\text{subject to:} \\
\quad k/100 \geq \alpha_i \text{ if } s_i^k \geq b_i, \quad i = 1, 2, \cdots, n \\
\quad 1 \leq x_i \leq n, \quad i = 1, 2, \cdots, n \\
\quad x_i \neq x_j, \quad i \neq j, \ i, j = 1, 2, \cdots, n \\
\quad 0 \leq y_1 \leq y_2 \leq \cdots \leq y_{m-1} \leq n \\
\quad x_i, y_j, \quad i = 1, 2, \cdots, n, \quad j = 1, 2, \cdots, m-1, \quad \text{integers.}
\end{cases}
$$

(2.122)

Numerical Experiment

Assume that there are 7 customers with the following time windows,

Node	Window	Node	Window
1	$[7:00, 9:00]$	5	$[15:00, 17:00]$
2	$[7:00, 9:00]$	6	$[19:00, 21:00]$
3	$[15:00, 17:00]$	7	$[19:00, 21:00]$
4	$[15:00, 17:00]$		

and each customer is visited within time windows with confidence level 0.90. We also assume that the distances are

$$D_{ij} = |i - j|, \quad i, j = 0, 1, 2, \cdots, 7$$

and travel times are normal uncertain variables

$$T_{ij} \sim \mathcal{N}(2|i - j|, 1), \quad i, j = 0, 1, 2, \cdots, 7.$$

A run of computer program (http://orsc.edu.cn/liu/resources.htm) shows that the optimal operational plan is

Vehicle 1: depot$\rightarrow 1 \rightarrow 3 \rightarrow$depot, starting time: 6:18

Vehicle 2: deport$\rightarrow 2 \rightarrow 5 \rightarrow 7 \rightarrow$depot, starting time: 4:18
Vehicle 3: depot$\rightarrow 4 \rightarrow 6 \rightarrow$depot, starting time: 8:18

whose total travel distance is 32.

2.10   Facility Location Problem

Facility location problem is to find locations for new facilities such that the conveying cost from facilities to customers is minimized. Facility location problem has been studied for half a century because of its widely practical application backgrounds. In practice, some factors such as demands, allocations, even locations of customers and facilities are usually changing. Liu [119] first introduced uncertainty theory into facility location problem and constructed some facility location models via uncertain programming.

In order to model facility location problem, we use the following indices, parameters, and decision variables:

$i = 1, 2, \cdots, n$: facilities;

$j = 1, 2, \cdots, m$: customers;

(a_j, b_j): location of customer j, $1 \leq j \leq m$;

ξ_j: uncertain demand of customer j, $1 \leq j \leq m$;

Φ_j: uncertainty distribution of ξ_j;

s_i: capacity of facility i, $1 \leq i \leq n$;

(x_i, y_i): decision vector representing the location of facility i, $1 \leq i \leq n$;

z_{ij}: quantity supplied to customer j by facility i after the uncertain demands are realized, $1 \leq i \leq n$, $1 \leq j \leq m$.

Conveying Cost

Write $\boldsymbol{\xi} = (\xi_1, \xi_2, \cdots, \xi_m)$. Each uncertain demand ξ_j of customer j ($1 \leq j \leq m$) is represented by the 99 method as

0.01	0.02	0.03	$\cdots$	0.99
d_j^1	d_j^2	d_j^3	$\cdots$	d_j^{99}

$\qquad(2.123)$

At each level $\alpha = k/100$ ($1 \leq k \leq 99$), the demand vector $\boldsymbol{\xi}$ becomes

$$\boldsymbol{d}_k = (d_1^k, d_2^k, \cdots, d_m^k). \qquad (2.124)$$

Note that the customers may be supplied by multiple facilities, rather than the nearest facility only. The feasible allocation set of $\boldsymbol{z}$ corresponding to $\boldsymbol{d}_k$ is represented by

$$Z_k = \left\{ \boldsymbol{z} \mid \begin{array}{l} z_{ij} \geq 0, \ i = 1, 2, \cdots, n, \ j = 1, 2, \cdots, m \\ \sum\limits_{i=1}^{n} z_{ij} = d_j^k, \ j = 1, 2, \cdots, m \\ \sum\limits_{j=1}^{m} z_{ij} \leq s_i, \ i = 1, 2, \cdots, n \end{array} \right\}. \qquad (2.125)$$

Note that Z_k may be an empty set for some k. The minimal conveying cost from facilities to customers is

$$C(\boldsymbol{x}, \boldsymbol{y}|k) = \min_{\boldsymbol{z} \in Z_k} \sum_{i=1}^{n} \sum_{j=1}^{m} z_{ij} \sqrt{(x_i - a_j)^2 + (y_i - b_j)^2} \qquad (2.126)$$

whose optimal solution $\boldsymbol{z}^*$ is called the optimal allocation corresponding to $\boldsymbol{d}_k$. If $Z_k = \emptyset$, then the demands of some customers are impossible to meet. For this case, as a penalty, we set

$$C(\boldsymbol{x}, \boldsymbol{y}|k) = \sum_{j=1}^{m} \max_{1 \le i \le n} d_j^k \sqrt{(x_i - a_j)^2 + (y_i - b_j)^2}. \qquad (2.127)$$

Since the conveying cost $C(\boldsymbol{x}, \boldsymbol{y}, \boldsymbol{\xi})$ is an increasing function with respect to $\boldsymbol{\xi}$, it is produced by the 99 method as

0.01	0.02	0.03	$\cdots$	0.99
$C(\boldsymbol{x}, \boldsymbol{y}|1)$	$C(\boldsymbol{x}, \boldsymbol{y}|2)$	$C(\boldsymbol{x}, \boldsymbol{y}|3)$	$\cdots$	$C(\boldsymbol{x}, \boldsymbol{y}|99)$

$$(2.128)$$

where $C(\boldsymbol{x}, \boldsymbol{y}|k)$ are determined by (2.126) and (2.127) for $k = 1, 2, \cdots, 99$. Note that the expected conveying cost may be approximately determined by

$$E[C(\boldsymbol{x}, \boldsymbol{y}, \boldsymbol{\xi})] = (C(\boldsymbol{x}, \boldsymbol{y}|1) + C(\boldsymbol{x}, \boldsymbol{y}|2) + \cdots + C(\boldsymbol{x}, \boldsymbol{y}|99))/99. \qquad (2.129)$$

Facility Location Model

In order to minimize the expected conveying cost, we have the following facility location model,

$$\begin{cases} \min_{\boldsymbol{x}, \boldsymbol{y}} E[C(\boldsymbol{x}, \boldsymbol{y}, \boldsymbol{\xi})] \\ \text{subject to:} \\ \quad g_j(\boldsymbol{x}, \boldsymbol{y}) \le 0, \ j = 1, 2, \cdots, p \end{cases} \qquad (2.130)$$

where $g_j(\boldsymbol{x}, \boldsymbol{y}) \le 0, \ j = 1, 2, \cdots, p$ represent the potential region of locations of new facilities. By using (2.129), the facility location model is simplified as follows,

$$\begin{cases} \min_{\boldsymbol{x}, \boldsymbol{y}} (C(\boldsymbol{x}, \boldsymbol{y}|1) + C(\boldsymbol{x}, \boldsymbol{y}|2) + \cdots + C(\boldsymbol{x}, \boldsymbol{y}|99))/99 \\ \text{subject to:} \\ \quad g_j(\boldsymbol{x}, \boldsymbol{y}) \le 0, \ j = 1, 2, \cdots, p. \end{cases} \qquad (2.131)$$

2.11 Machine Scheduling Problem

Machine scheduling problem is concerned with finding an efficient schedule during an uninterrupted period of time for a set of machines to process a

set of jobs. Much of research work has been done on this type of problem during the past five decades. The study of machine scheduling problem with uncertain processing times was started by Liu [129] in 2009.

In a machine scheduling problem, we assume that (a) each job can be processed on any machine without interruption; and (b) each machine can process only one job at a time. We also use the following indices and parameters:

$i = 1, 2, \cdots, n$: jobs;

$k = 1, 2, \cdots, m$: machines;

ξ_{ik}: uncertain processing time of job i on machine k;

Φ_{ik}: uncertainty distribution of ξ_{ik}.

How to Represent a Schedule?

The schedule is represented by Liu's formulation [119] via two decision vectors $\boldsymbol{x}$ and $\boldsymbol{y}$, where

$\boldsymbol{x} = (x_1, x_2, \cdots, x_n)$: integer decision vector representing n jobs with $1 \leq x_i \leq n$ and $x_i \neq x_j$ for all $i \neq j$, $i, j = 1, 2, \cdots, n$. That is, the sequence $\{x_1, x_2, \cdots, x_n\}$ is a rearrangement of $\{1, 2, \cdots, n\}$;

$\boldsymbol{y} = (y_1, y_2, \cdots, y_{m-1})$: integer decision vector with $y_0 \equiv 0 \leq y_1 \leq y_2 \leq \cdots \leq y_{m-1} \leq n \equiv y_m$.

We note that the schedule is fully determined by the decision vectors $\boldsymbol{x}$ and $\boldsymbol{y}$ in the following way. For each k $(1 \leq k \leq m)$, if $y_k = y_{k-1}$, then the machine k is not used; if $y_k > y_{k-1}$, then the machine k is used and processes jobs $x_{y_{k-1}+1}, x_{y_{k-1}+2}, \cdots, x_{y_k}$ in turn. Thus the schedule of all machines is as follows,

$$
\begin{aligned}
&\text{Machine 1: } x_{y_0+1} \rightarrow x_{y_0+2} \rightarrow \cdots \rightarrow x_{y_1}; \\
&\text{Machine 2: } x_{y_1+1} \rightarrow x_{y_1+2} \rightarrow \cdots \rightarrow x_{y_2}; \\
&\qquad \cdots \\
&\text{Machine } m: x_{y_{m-1}+1} \rightarrow x_{y_{m-1}+2} \rightarrow \cdots \rightarrow x_{y_m}.
\end{aligned}
\tag{2.132}
$$

Completion Times

Let $C_i(\boldsymbol{x}, \boldsymbol{y}, \boldsymbol{\xi})$ be the completion times of jobs i, $i = 1, 2, \cdots, n$, respectively. For each k with $1 \leq k \leq m$, if the machine k is used (i.e., $y_k > y_{k-1}$), then we have

$$
C_{x_{y_{k-1}+1}}(\boldsymbol{x}, \boldsymbol{y}, \boldsymbol{\xi}) = \xi_{x_{y_{k-1}+1}k}
\tag{2.133}
$$

and

$$
C_{x_{y_{k-1}+j}}(\boldsymbol{x}, \boldsymbol{y}, \boldsymbol{\xi}) = C_{x_{y_{k-1}+j-1}}(\boldsymbol{x}, \boldsymbol{y}, \boldsymbol{\xi}) + \xi_{x_{y_{k-1}+j}k}
\tag{2.134}
$$

for $2 \leq j \leq y_k - y_{k-1}$.

Assume that each uncertain processing time ξ_{ik} of job i on machine j is represented by the 99 method as

$$
\begin{array}{|c|c|c|c|c|}
\hline
0.01 & 0.02 & 0.03 & \cdots & 0.99 \\
\hline
t_{ik}^1 & t_{ik}^2 & t_{ik}^3 & \cdots & t_{ik}^{99} \\
\hline
\end{array}
\tag{2.135}
$$

If the machine k is used, then the completion time $C_{x_{y_{k-1}+1}}(\boldsymbol{x}, \boldsymbol{y}, \boldsymbol{\xi})$ of job $x_{y_{k-1}+1}$ is an uncertain variable whose uncertainty distribution is

$$
\Psi_{x_{y_{k-1}+1}}(z) = \Phi_{x_{y_{k-1}+1}k}(z)
\tag{2.136}
$$

and is produced by the 99 method as

$$
\begin{array}{|c|c|c|c|}
\hline
0.01 & 0.02 & \cdots & 0.99 \\
\hline
s_{x_{y_{k-1}+1}}^1 & s_{x_{y_{k-1}+1}}^2 & \cdots & s_{x_{y_{k-1}+1}}^{99} \\
\hline
\end{array}
\tag{2.137}
$$

Generally, suppose that the completion time $C_{x_{y_{k-1}+j-1}}(\boldsymbol{x}, \boldsymbol{y}, \boldsymbol{\xi})$ has an uncertainty distribution $\Psi_{x_{y_{k-1}+j-1}}(z)$ and is produced by the 99 method as

$$
\begin{array}{|c|c|c|c|}
\hline
0.01 & 0.02 & \cdots & 0.99 \\
\hline
s_{x_{y_{k-1}+j-1}}^1 & s_{x_{y_{k-1}+j-1}}^2 & \cdots & s_{x_{y_{k-1}+j-1}}^{99} \\
\hline
\end{array}
\tag{2.138}
$$

Then the completion time $C_{x_{y_{k-1}+j}}(\boldsymbol{x}, \boldsymbol{y}, \boldsymbol{\xi})$ has an uncertainty distribution

$$
\Psi_{x_{y_{k-1}+j}}(z) = \sup_{z_1+z_2=z} \Psi_{x_{y_{k-1}+j-1}}(z_1) \wedge \Phi_{x_{y_{k-1}+j-1}k}(z_2)
\tag{2.139}
$$

and is produced by the 99 method as

$$
\begin{array}{|c|c|c|}
\hline
0.01 & \cdots & 0.99 \\
\hline
s_{x_{y_{k-1}+j-1}}^1 + t_{x_{y_{k-1}+j-1}k}^1 & \cdots & s_{x_{y_{k-1}+j-1}}^{99} + t_{x_{y_{k-1}+j-1}k}^{99} \\
\hline
\end{array}
$$

where $s_{x_{y_{k-1}+j-1}}^1, s_{x_{y_{k-1}+j-1}}^2, \cdots, s_{x_{y_{k-1}+j-1}}^{99}$ are determined by (2.138), and $t_{x_{y_{k-1}+j-1}k}^1, t_{x_{y_{k-1}+j-1}k}^2, \cdots, t_{x_{y_{k-1}+j-1}k}^{99}$ are determined by (2.135). This recursive process may produce all completion times of jobs.

Makespan

Note that, for each k ($1 \leq k \leq m$), the value $C_{x_{y_k}}(\boldsymbol{x}, \boldsymbol{y}, \boldsymbol{\xi})$ is just the time that the machine k finishes all jobs assigned to it, and is produced by the 99 method as

$$
\begin{array}{|c|c|c|c|}
\hline
0.01 & 0.02 & \cdots & 0.99 \\
\hline
s_{x_{y_k}}^1 & s_{x_{y_k}}^2 & \cdots & s_{x_{y_k}}^{99} \\
\hline
\end{array}
\tag{2.140}
$$

Thus the makespan of the schedule $(\boldsymbol{x}, \boldsymbol{y})$ is determined by

$$
f(\boldsymbol{x}, \boldsymbol{y}, \boldsymbol{\xi}) = \max_{1 \leq k \leq m} C_{x_{y_k}}(\boldsymbol{x}, \boldsymbol{y}, \boldsymbol{\xi})
\tag{2.141}
$$

whose uncertainty distribution is

$$\Upsilon(z) = \min_{1 \le k \le m} \Psi_{x_{y_k}}(z) \qquad (2.142)$$

and is produced by the 99 method as

0.01	0.02	$\cdots$	0.99
$\displaystyle\bigvee_{k=1}^{m} s^1_{x_{y_k}}$	$\displaystyle\bigvee_{k=1}^{m} s^2_{x_{y_k}}$	$\cdots$	$\displaystyle\bigvee_{k=1}^{m} s^{99}_{x_{y_k}}$

$$(2.143)$$

Machine Scheduling Model

In order to minimize the expected makespan $E[f(\boldsymbol{x}, \boldsymbol{y}, \boldsymbol{\xi})]$, we have the following machine scheduling model,

$$\begin{cases} \min E[f(\boldsymbol{x}, \boldsymbol{y}, \boldsymbol{\xi})] \\ \text{subject to:} \\ \quad 1 \le x_i \le n, \quad i = 1, 2, \cdots, n \\ \quad x_i \ne x_j, \quad i \ne j, \ i, j = 1, 2, \cdots, n \\ \quad 0 \le y_1 \le y_2 \cdots \le y_{m-1} \le n \\ \quad x_i, y_j, \quad i = 1, 2, \cdots, n, \quad j = 1, 2, \cdots, m-1, \quad \text{integers.} \end{cases} \qquad (2.144)$$

By using (2.143), the machine scheduling model is simplified as follows,

$$\begin{cases} \min \left(\displaystyle\bigvee_{k=1}^{m} s^1_{x_{y_k}} + \bigvee_{k=1}^{m} s^2_{x_{y_k}} + \cdots + \bigvee_{k=1}^{m} s^{99}_{x_{y_k}} \right) / 99 \\ \text{subject to:} \\ \quad 1 \le x_i \le n, \quad i = 1, 2, \cdots, n \\ \quad x_i \ne x_j, \quad i \ne j, \ i, j = 1, 2, \cdots, n \\ \quad 0 \le y_1 \le y_2 \cdots \le y_{m-1} \le n \\ \quad x_i, y_j, \quad i = 1, 2, \cdots, n, \quad j = 1, 2, \cdots, m-1, \quad \text{integers.} \end{cases} \qquad (2.145)$$

Chapter 3

Uncertain Process

An uncertain process is essentially a sequence of uncertain variables indexed by time or space. The study of uncertain process was started by Liu [128] in 2008. This chapter introduces the basic concepts of uncertain process, including renewal process, martingale, Markov process and stationary process.

3.1 Definitions

Definition 3.1 *(Liu [128]) Let T be an index set and let $(\Gamma, \mathcal{L}, \mathcal{M})$ be an uncertainty space. An uncertain process is a measurable function from $T \times (\Gamma, \mathcal{L}, \mathcal{M})$ to the set of real numbers, i.e., for each $t \in T$ and any Borel set B of real numbers, the set*

$$\{X_t \in B\} = \{\gamma \in \Gamma \mid X_t(\gamma) \in B\} \tag{3.1}$$

is an event.

That is, an uncertain process $X_t(\gamma)$ is a function of two variables such that the function $X_{t^*}(\gamma)$ is an uncertain variable for each t^*.

Definition 3.2 *For each fixed γ^*, the function $X_t(\gamma^*)$ is called a sample path of the uncertain process X_t.*

Definition 3.3 *An uncertain process X_t is said to be sample-continuous if almost all sample paths are continuous with respect to t.*

Definition 3.4 *An uncertain process X_t is said to have independent increments if*

$$X_{t_1} - X_{t_0}, \ X_{t_2} - X_{t_1}, \ \cdots, \ X_{t_k} - X_{t_{k-1}} \tag{3.2}$$

are independent uncertain variables for any times $t_0 < t_1 < \cdots < t_k$.

Definition 3.5 *An uncertain process X_t is said to have stationary increments if, for any given $t > 0$, the increments $X_{s+t} - X_s$ are identically distributed uncertain variables for all $s > 0$.*

Definition 3.6 *For any partition of closed interval $[0, t]$ with $0 = t_1 < t_2 < \cdots < t_{k+1} = t$, the mesh is written as*

$$\Delta = \max_{1 \leq i \leq k} |t_{i+1} - t_i|.$$

Let $m > 0$ be a real number. Then the m-variation of uncertain process X_t is

$$\|X\|_t^m = \lim_{\Delta \to 0} \sum_{i=1}^{k} |X_{t_{i+1}} - X_{t_i}|^m \tag{3.3}$$

provided that the limit exists almost surely and is an uncertain process. Especially,

$$\|X\|_t = \lim_{\Delta \to 0} \sum_{i=1}^{k} |X_{t_{i+1}} - X_{t_i}| \tag{3.4}$$

is called total variation, and

$$\|X\|_t^2 = \lim_{\Delta \to 0} \sum_{i=1}^{k} |X_{t_{i+1}} - X_{t_i}|^2 \tag{3.5}$$

is called the squared variation of uncertain process X_t.

3.2   Renewal Process

Definition 3.7 *(Liu [128]) Let $\xi_1, \xi_2, \cdots$ be iid positive uncertain variables. Define $S_0 = 0$ and $S_n = \xi_1 + \xi_2 + \cdots + \xi_n$ for $n \geq 1$. Then the uncertain process*

$$N_t = \max_{n \geq 0} \left\{ n \mid S_n \leq t \right\} \tag{3.6}$$

is called a renewal process.

If $\xi_1, \xi_2, \cdots$ denote the interarrival times of successive events. Then S_n can be regarded as the waiting time until the occurrence of the nth event, and N_t is the number of renewals in $(0, t]$. The renewal process N_t is not sample-continuous. But each sample path of N_t is a right-continuous and increasing step function taking only nonnegative integer values. Furthermore, the size of each jump of N_t is always 1. In other words, N_t has at most one renewal at each time. In particular, N_t does not jump at time 0. Since $N_t \geq n$ is equivalent to $S_n \leq t$, we immediately have

$$\mathcal{M}\{N_t \geq n\} = \mathcal{M}\{S_n \leq t\}. \tag{3.7}$$
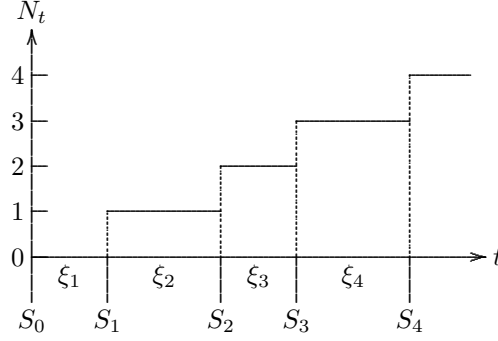
Figure 3.1: A Sample Path of Renewal Process

Since $N_t \leq n$ is equivalent to $S_{n+1} > t$, by using the self-duality axiom, we immediately have

$$\mathcal{M}\{N_t \leq n\} = 1 - \mathcal{M}\{S_{n+1} \leq t\}. \tag{3.8}$$

Theorem 3.1 *Let N_t be a renewal process with uncertain interarrival times $\xi_1, \xi_2, \cdots$ If those interarrival times have a common uncertainty distribution Φ, then N_t has an uncertainty distribution*

$$\Upsilon_t(x) = 1 - \Phi\left(\frac{t}{\lfloor x \rfloor + 1}\right), \quad \forall x \geq 0 \tag{3.9}$$

where $\lfloor x \rfloor$ represents the maximal integer less than or equal to x.

Proof: Note that S_{n+1} has an uncertainty distribution $\Phi(x/(n+1))$. It follows from (3.8) that

$$\mathcal{M}\{N_t \leq n\} = 1 - \mathcal{M}\{S_{n+1} \leq t\} = 1 - \Phi\left(\frac{t}{n+1}\right).$$

Since N_t takes integer values, for any $x \geq 0$, we have

$$\Upsilon_t(x) = \mathcal{M}\{N_t \leq x\} = \mathcal{M}\{N_t \leq \lfloor x \rfloor\} = 1 - \Phi\left(\frac{t}{\lfloor x \rfloor + 1}\right).$$

The theorem is verified.

Theorem 3.2 *Let N_t be a renewal process with uncertain interarrival times $\xi_1, \xi_2, \cdots$ If those interarrival times have a common uncertainty distribution Φ, then*

$$E[N_t] = \sum_{n=1}^{\infty} \Phi\left(\frac{t}{n}\right). \tag{3.10}$$
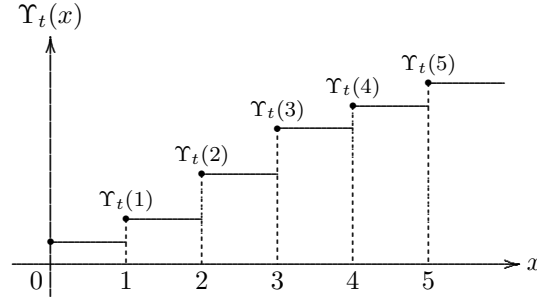
Figure 3.2: Uncertainty Distribution Υ_t of Renewal Process N_t

Proof: Since N_t takes only nonnegative integer values, it follows from the definition of expected value and Theorem 3.1 that

$$E[N_t] = \int_0^\infty (1 - \Upsilon_t(x))\mathrm{d}x = \sum_{n=0}^\infty (1 - \Upsilon_t(n))$$

$$= \sum_{n=0}^\infty \Phi\left(\frac{t}{n+1}\right) = \sum_{n=1}^\infty \Phi\left(\frac{t}{n}\right).$$

Thus the theorem is verified.

Theorem 3.3 *(Renewal Theorem) Let N_t be a renewal process with uncertain interarrival times $\xi_1, \xi_2, \cdots$ Then*

$$\lim_{t\to\infty} \frac{E[N_t]}{t} = E\left[\frac{1}{\xi_1}\right]. \tag{3.11}$$

If those interarrival times have a common uncertainty distribution Φ, then

$$\lim_{t\to\infty} \frac{E[N_t]}{t} = \int_0^{+\infty} \Phi\left(\frac{1}{x}\right)\mathrm{d}x. \tag{3.12}$$

If the uncertainty distribution Φ is regular, then

$$\lim_{t\to\infty} \frac{E[N_t]}{t} = \int_0^1 \frac{1}{\Phi^{-1}(\alpha)}\mathrm{d}\alpha. \tag{3.13}$$

Proof: The uncertainty distribution Υ_t of N_t has been given by Theorem 3.1. It follows from the operational law that the uncertainty distribution of N_t/t is

$$\Psi_t(x) = \Upsilon_t(tx) = 1 - \Phi\left(\frac{t}{\lfloor tx\rfloor + 1}\right)$$

where $\lfloor tx\rfloor$ represents the maximal integer less than or equal to $\lfloor tx\rfloor$. Thus

$$\frac{E[N_t]}{t} = \int_0^{+\infty} (1 - \Psi_t(x))\mathrm{d}x.$$

On the other hand, $1/\xi_1$ has an uncertainty distribution $1 - \Phi(1/x)$ whose expected value is

$$E\left[\frac{1}{\xi_1}\right] = \int_0^{+\infty} \Phi\left(\frac{1}{x}\right) \mathrm{d}x.$$

Note that

$$(1 - \Psi_t(x)) \le \Phi\left(\frac{1}{x}\right), \quad \forall t, x$$

and

$$\lim_{t\to\infty} (1 - \Psi_t(x)) = \Phi\left(\frac{1}{x}\right), \quad \forall x.$$

It follows from Lebesgue dominated convergence theorem that

$$\lim_{t\to\infty} \frac{E[N_t]}{t} = \lim_{t\to\infty} \int_0^{+\infty} (1 - \Psi_t(x))\mathrm{d}x = \int_0^{+\infty} \Phi\left(\frac{1}{x}\right)\mathrm{d}x = E\left[\frac{1}{\xi_1}\right].$$

Furthermore, since the uncertainty distribution of $1/\xi$ has an inverse function $1/\Phi^{-1}(1 - \alpha)$, we get

$$E\left[\frac{1}{\xi}\right] = \int_0^1 \frac{1}{\Phi^{-1}(1 - \alpha)}\mathrm{d}\alpha = \int_0^1 \frac{1}{\Phi^{-1}(\alpha)}\mathrm{d}\alpha$$

The theorem is proved.
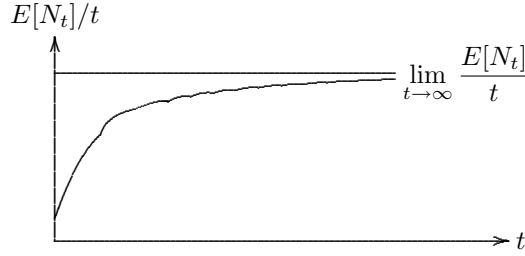
$E[N_t]/t$

Figure 3.3: Average Renewal Number $E[N_t]/t$

Example 3.1: A renewal process N_t is called a *linear renewal process* if $\xi_1, \xi_2, \cdots$ are iid linear uncertain variables $\mathcal{L}(a, b)$ with $a > 0$. It follows from the renewal theorem that

$$\lim_{t\to\infty} \frac{E[N_t]}{t} = \frac{\ln b - \ln a}{b - a}. \tag{3.14}$$

Example 3.2: A renewal process N_t is called a *zigzag renewal process* if $\xi_1, \xi_2, \cdots$ are iid zigzag uncertain variables $\mathcal{Z}(a, b, c)$ with $a > 0$. It follows from the renewal theorem that

$$\lim_{t\to\infty} \frac{E[N_t]}{t} = \frac{1}{2}\left(\frac{\ln b - \ln a}{b - a} + \frac{\ln c - \ln b}{c - b}\right). \tag{3.15}$$

Example 3.3: A renewal process N_t is called a *lognormal renewal process* if $\xi_1, \xi_2, \cdots$ are iid lognormal uncertain variables $\mathcal{LOGN}(e, \sigma)$. If $\sigma < \pi/\sqrt{3}$, then

$$\lim_{t \to \infty} \frac{E[N_t]}{t} = \sqrt{3}\sigma \exp(-e) \csc(\sqrt{3}\sigma). \tag{3.16}$$

Otherwise, we have

$$\lim_{t \to \infty} \frac{E[N_t]}{t} = +\infty. \tag{3.17}$$

3.3   Martingale

Definition 3.8 *An uncertain process X_t is called martingale if it has independent increments whose expected values are zero.*

3.4   Markov Process

Definition 3.9 *An uncertain process X_t is called Markov if, given the value of X_t, the uncertain variables X_s and X_u are independent for any $s > t > u$.*

3.5   Stationary Process

Definition 3.10 *An uncertain process X_t is called stationary if for any positive integer k and any times $t_1, t_2, \cdots, t_k$ and s, the uncertain vectors*

$$(X_{t_1}, X_{t_2}, \cdots, X_{t_k}) \quad and \quad (X_{t_1+s}, X_{t_2+s}, \cdots, X_{t_k+s}) \tag{3.18}$$

are identically distributed.

Chapter 4

Uncertain Calculus

Uncertain calculus, invented by Liu [130] in 2009, is a branch of mathematics that deals with differentiation and integration of function of uncertain processes. We will introduce canonical process, uncertain integral, chain rule, and integration by parts.

4.1 Canonical Process

Definition 4.1 *(Liu [130]) An uncertain process C_t is said to be a canonical process if*
(i) $C_0 = 0$ and almost all sample paths are Lipschitz continuous,
(ii) C_t has stationary and independent increments,
(iii) every increment $C_{s+t} - C_s$ is a normal uncertain variable with expected value 0 and variance t^2.
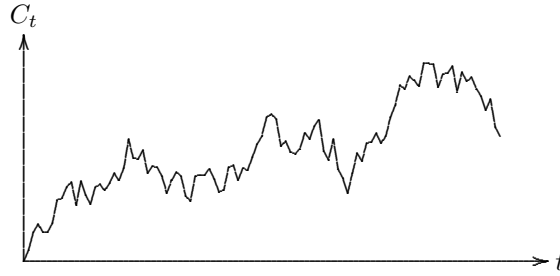
Figure 4.1: A Sample Path of Canonical Process

Note that almost all sample paths of canonical process are Lipschitz continuous functions, but almost all sample paths of Brownian motion are continuous but non-Lipschitz functions. If we say Brownian motion describes

the irregular movement of pollen with infinite speed, then we may say the canonical process describes the irregular movement of pollen with finite speed.

Theorem 4.1 *(Existence Theorem) There is a canonical process.*

Proof: Without loss of generality, we only prove that there is a canonical process on the range of $t \in [0, 1]$. Let

$$\big\{ \xi(r) \,\big|\, r \text{ represents rational numbers in } [0, 1] \big\}$$

be a countable sequence of independently normal uncertain variables with expected value zero and variance one. For each integer n, we define an uncertain process

$$X_n(t) = \begin{cases} \dfrac{1}{n} \displaystyle\sum_{i=1}^{k} \xi\left(\dfrac{i}{n}\right), & \text{if } t = \dfrac{k}{n} \quad (k = 0, 1, \cdots, n) \\[2mm] \qquad \text{linear}, & \text{otherwise.} \end{cases}$$

Since the limit

$$\lim_{n \to \infty} X_n(t)$$

exists almost surely, we may verify that the limit meets the conditions of canonical process. Hence there is a canonical process.

Theorem 4.2 *Let C_t be a canonical process. Then for each time $t > 0$, the ratio C_t/t is a normal uncertain variable with expected value 0 and variance 1. That is,*

$$\frac{C_t}{t} \sim \mathcal{N}(0, 1) \tag{4.1}$$

for any $t > 0$.

Proof: It follows from the definition of canonical process that at each time t, C_t is a normal uncertain variable with uncertainty distribution

$$\Phi(x) = \left(1 + \exp\left(-\frac{\pi x}{\sqrt{3}t} \right) \right)^{-1}.$$

Thus C_t/t has an uncertainty distribution

$$\Psi(x) = \Phi(tx) = \left(1 + \exp\left(-\frac{\pi x}{\sqrt{3}} \right) \right)^{-1}.$$

Hence C_t/t is a normal uncertain variable with expected value 0 and variance 1. The theorem is verified.

Theorem 4.3 *Let C_t be a canonical process. Then for any level $x \in \Re$ and any time $t > 0$, we have*

$$\mathcal{M}\{C_t \le x\} = \left(1 + \exp\left(-\frac{\pi x}{\sqrt{3}t}\right)\right)^{-1}, \qquad (4.2)$$

$$\mathcal{M}\{C_t \ge x\} = \left(1 + \exp\left(\frac{\pi x}{\sqrt{3}t}\right)\right)^{-1}. \qquad (4.3)$$

Proof: Since C_t is a normal uncertain variable with expected value 0 and variance t^2, we get (4.2) immediately. The equation (4.3) may be derived from $\mathcal{M}\{C_t \ge x\} = 1 - \mathcal{M}\{C_t \le x\}$.

Arithmetic Canonical Process

Definition 4.2 *Let C_t be a canonical process. Then for any real numbers e and σ,*

$$A_t = et + \sigma C_t \qquad (4.4)$$

is called an arithmetic canonical process, where e is called the drift and σ is called the diffusion.

At each time t, A_t is a normal uncertain variable $\mathcal{N}(et, \sigma t)$. That is, the expected value $E[A_t] = et$ and variance $V[A_t] = \sigma^2 t^2$ at any time t.

Geometric Canonical Process

Definition 4.3 *Let C_t be a canonical process. Then for any real numbers e and σ,*

$$G_t = \exp(et + \sigma C_t) \qquad (4.5)$$

is called a geometric canonical process, where e is called the log-drift and σ is called the log-diffusion.

At each time t, G_t is a lognormal uncertain variable $\mathcal{LOGN}(et, \sigma t)$. If $t < \pi/(\sigma\sqrt{3})$, then $E[G_t] = \sqrt{3}\sigma t \exp(et) \csc(\sqrt{3}\sigma t)$. However, when t arrives at $\pi/(\sigma\sqrt{3})$, we have $E[G_t] = +\infty$.

4.2   Uncertain Integral

Definition 4.4 *(Liu [130]) Let X_t be an uncertain process and let C_t be a canonical process. For any partition of closed interval $[a, b]$ with $a = t_1 < t_2 < \cdots < t_{k+1} = b$, the mesh is written as*

$$\Delta = \max_{1 \le i \le k} |t_{i+1} - t_i|. \qquad (4.6)$$

Then the uncertain integral of X_t with respect to C_t is

$$\int_a^b X_t \mathrm{d}C_t = \lim_{\Delta \to 0} \sum_{i=1}^k X_{t_i} \cdot (C_{t_{i+1}} - C_{t_i}) \tag{4.7}$$

provided that the limit exists almost surely and is an uncertain variable.

Example 4.1: Let C_t be a canonical process. Then for any partition $0 = t_1 < t_2 < \cdots < t_{k+1} = s$, we have

$$\int_0^s \mathrm{d}C_t = \lim_{\Delta \to 0} \sum_{i=1}^k (C_{t_{i+1}} - C_{t_i}) \equiv C_s - C_0 = C_s.$$

Example 4.2: Let C_t be a canonical process. Then for any partition $0 = t_1 < t_2 < \cdots < t_{k+1} = s$, we have

$$C_s^2 = \sum_{i=1}^k \left(C_{t_{i+1}}^2 - C_{t_i}^2 \right)$$

$$= \sum_{i=1}^k \left(C_{t_{i+1}} - C_{t_i} \right)^2 + 2 \sum_{i=1}^k C_{t_i} \left(C_{t_{i+1}} - C_{t_i} \right)$$

$$\to 0 + 2 \int_0^s C_t \mathrm{d}C_t$$

as $\Delta \to 0$. That is,

$$\int_0^s C_t \mathrm{d}C_t = \frac{1}{2} C_s^2.$$

Example 4.3: Let C_t be a canonical process. Then for any number α $(0 < \alpha < 1)$, the uncertain process

$$F_s = \int_0^s (s - t)^{-\alpha} \mathrm{d}C_t \tag{4.8}$$

is called a *fractional canonical process* with index α. At each time s, F_s is a normal uncertain variable with expected value 0 and variance

$$V[F_s] = \frac{s^{2-2\alpha}}{(1-\alpha)^2}. \tag{4.9}$$

Example 4.4: Generally, let C_t be a canonical process and let $f(t)$ be an integrable function with respect to t. Then the uncertain integral

$$\int_0^s f(t) \mathrm{d}C_t \tag{4.10}$$

is a normal uncertain variable with expected value 0 and variance

$$\left(\int_0^s |f(t)|\mathrm{d}t\right)^2. \tag{4.11}$$

4.3   Chain Rule

Theorem 4.4 *(Liu [130]) Let C_t be a canonical process, and let $h(t,c)$ be a continuously differentiable function. Define $X_t = h(t, C_t)$. Then we have the following chain rule*

$$\mathrm{d}X_t = \frac{\partial h}{\partial t}(t, C_t)\mathrm{d}t + \frac{\partial h}{\partial c}(t, C_t)\mathrm{d}C_t. \tag{4.12}$$

Proof: Write $\Delta C_t = C_{t+\Delta t} - C_t = C_{\Delta t}$. Then Δt and ΔC_t are infinitesimals with the same order. Since the function h is continuously differentiable, by using Taylor series expansion, the infinitesimal increment of X_t has a first-order approximation

$$\Delta X_t = \frac{\partial h}{\partial t}(t, C_t)\Delta t + \frac{\partial h}{\partial c}(t, C_t)\Delta C_t.$$

Hence we obtain the chain rule because it makes

$$X_s = X_0 + \int_0^s \frac{\partial h}{\partial t}(t, C_t)\mathrm{d}t + \int_0^s \frac{\partial h}{\partial c}(t, C_t)\mathrm{d}C_t$$

for any $s \geq 0$.

Remark 4.1: The infinitesimal increment $\mathrm{d}C_t$ in (4.12) may be replaced with the derived canonical process

$$\mathrm{d}Y_t = u_t\mathrm{d}t + v_t\mathrm{d}C_t \tag{4.13}$$

where u_t and v_t are absolutely integrable uncertain processes, thus producing

$$\mathrm{d}h(t, Y_t) = \frac{\partial h}{\partial t}(t, Y_t)\mathrm{d}t + \frac{\partial h}{\partial c}(t, Y_t)\mathrm{d}Y_t. \tag{4.14}$$

Example 4.5: Applying the chain rule, we obtain the following formula

$$\mathrm{d}(tC_t) = C_t\mathrm{d}t + t\mathrm{d}C_t.$$

Hence we have

$$sC_s = \int_0^s \mathrm{d}(tC_t) = \int_0^s C_t\mathrm{d}t + \int_0^s t\mathrm{d}C_t.$$

That is,

$$\int_0^s t\mathrm{d}C_t = sC_s - \int_0^s C_t\mathrm{d}t.$$

Example 4.6: Applying the chain rule, we obtain the following formula

$$\mathrm{d}(C_t^2) = 2C_t\mathrm{d}C_t.$$

Then we have

$$C_s^2 = \int_0^s \mathrm{d}(C_t^2) = 2\int_0^s C_t\mathrm{d}C_t.$$

It follows that

$$\int_0^s C_t\mathrm{d}C_t = \frac{1}{2}C_s^2.$$

Example 4.7: Applying the chain rule, we obtain the following formula

$$\mathrm{d}(C_t^3) = 3C_t^2\mathrm{d}C_t.$$

Thus we get

$$C_s^3 = \int_0^s \mathrm{d}(C_t^3) = 3\int_0^s C_t^2\mathrm{d}C_t.$$

That is

$$\int_0^s C_t^2\mathrm{d}C_t = \frac{1}{3}C_s^3.$$

4.4   Integration by Parts

Theorem 4.5 *(Integration by Parts) Suppose that C_t is a canonical process and $F(t)$ is an absolutely continuous function. Then*

$$\int_0^s F(t)\mathrm{d}C_t = F(s)C_s - \int_0^s C_t\mathrm{d}F(t). \tag{4.15}$$

Proof: By defining $h(t, C_t) = F(t)C_t$ and using the chain rule, we get

$$\mathrm{d}(F(t)C_t) = C_t\mathrm{d}F(t) + F(t)\mathrm{d}C_t.$$

Thus

$$F(s)C_s = \int_0^s \mathrm{d}(F(t)C_t) = \int_0^s C_t\mathrm{d}F(t) + \int_0^s F(t)\mathrm{d}C_t$$

which is just (4.15).

Example 4.8: Assume $F(t) \equiv 1$. Then by using the integration by parts, we immediately obtain

$$\int_0^s \mathrm{d}C_t = C_s.$$

Example 4.9: Assume $F(t) = t$. Then by using the integration by parts, we immediately obtain

$$\int_0^s t\mathrm{d}C_t = sC_s - \int_0^s C_t\mathrm{d}t.$$

Example 4.10: Assume $F(t) = t^2$. Then by using the integration by parts, we obtain

$$\int_0^s t^2 \mathrm{d}C_t = s^2 C_s - \int_0^s C_t \mathrm{d}t^2 = s^2 C_s - 2 \int_0^s t C_t \mathrm{d}t$$
$$= (s^2 - 2s)C_s + 2 \int_0^s C_t \mathrm{d}t.$$

Example 4.11: Assume $F(t) = \sin t$. Then by using the integration by parts, we obtain

$$\int_0^s \sin t \mathrm{d}C_t = C_s \sin s - \int_0^s C_t \mathrm{d}\sin t = C_s \sin s - \int_0^s C_t \cos t \mathrm{d}t.$$

Chapter 5

Uncertain Differential Equation

Uncertain differential equation, proposed by Liu [128] in 2008, is a type of differential equation driven by canonical process. This chapter will discuss the existence, uniqueness and stability of solutions of uncertain differential equations. This chapter also presents some applications of uncertain differential equation in finance, filtering, and control.

5.1 Uncertain Differential Equation

Definition 5.1 *(Liu [128]) Suppose C_t is a canonical process, and f and g are some given functions. Then*

$$\mathrm{d}X_t = f(t, X_t)\mathrm{d}t + g(t, X_t)\mathrm{d}C_t \tag{5.1}$$

is called an uncertain differential equation. A solution is an uncertain process X_t that satisfies (5.1) identically in t.

Remark 5.1: Note that there is no precise definition for the terms $\mathrm{d}X_t$, $\mathrm{d}t$ and $\mathrm{d}C_t$ in the uncertain differential equation (5.1). The mathematically meaningful form is the uncertain integral equation

$$X_s = X_0 + \int_0^s f(t, X_t)\mathrm{d}t + \int_0^s g(t, X_t)\mathrm{d}C_t. \tag{5.2}$$

However, the differential form is more convenient for us. This is the main reason why we accept the differential form.

Example 5.1: Let C_t be a canonical process. Then the uncertain differential equation

$$\mathrm{d}X_t = a\mathrm{d}t + b\mathrm{d}C_t \tag{5.3}$$

has a solution

$$X_t = at + bC_t \tag{5.4}$$

which is just an arithmetic canonical process.

Example 5.2: Let C_t be a canonical process. Then the uncertain differential equation

$$dX_t = aX_t dt + bX_t dC_t \tag{5.5}$$

has a solution

$$X_t = \exp(at + bC_t) \tag{5.6}$$

which is just a geometric canonical process.

Example 5.3: Let u_t and v_t be some continuous functions with respect to t. Consider the *homogeneous linear uncertain differential equation*

$$dX_t = u_t X_t dt + v_t X_t dC_t. \tag{5.7}$$

It follows from the chain rule that

$$d\ln X_t = \frac{dX_t}{X_t} = u_t dt + v_t dC_t.$$

Integration of both sides yields

$$\ln X_t - \ln X_0 = \int_0^t u_s ds + \int_0^t v_s dC_s.$$

Therefore the solution of (5.7) is

$$X_t = X_0 \exp\left(\int_0^t u_s ds + \int_0^t v_s dC_s\right). \tag{5.8}$$

Example 5.4: Suppose $u_{1t}, u_{2t}, v_{1t}, v_{2t}$ are continuous functions with respect to t. Consider the *linear uncertain differential equation*

$$dX_t = (u_{1t} X_t + u_{2t})dt + (v_{1t} X_t + v_{2t})dC_t. \tag{5.9}$$

At first, we define two uncertain processes U_t and V_t via

$$dU_t = u_{1t} U_t dt + v_{1t} U_t dC_t, \quad dV_t = \frac{u_{2t}}{U_t}dt + \frac{v_{2t}}{U_t}dC_t.$$

Then we have $X_t = U_t V_t$ because

$$\begin{aligned}
dX_t &= V_t dU_t + U_t dV_t \\
&= (u_{1t} U_t V_t + u_{2t})dt + (v_{1t} U_t V_t + v_{2t})dC_t \\
&= (u_{1t} X_t + u_{2t})dt + (v_{1t} X_t + v_{2t})dC_t.
\end{aligned}$$

Note that

$$U_t = U_0 \exp\left(\int_0^t u_{1s}\mathrm{d}s + \int_0^t v_{1s}\mathrm{d}C_s\right),$$

$$V_t = V_0 + \int_0^t \frac{u_{2s}}{U_s}\mathrm{d}s + \int_0^t \frac{v_{2s}}{U_s}\mathrm{d}C_s.$$

Taking $U_0 = 1$ and $V_0 = X_0$, we get a solution of the linear uncertain differential equation as follows,

$$X_t = U_t \exp\left(X_0 + \int_0^t \frac{u_{2s}}{U_s}\mathrm{d}s + \int_0^t \frac{v_{2s}}{U_s}\mathrm{d}C_s\right) \tag{5.10}$$

where

$$U_t = \exp\left(\int_0^t u_{1s}\mathrm{d}s + \int_0^t v_{1s}\mathrm{d}C_s\right). \tag{5.11}$$

Example 5.5: Let C_t be a canonical process. Then the special linear uncertain differential equation

$$\mathrm{d}X_t = (m - aX_t)\mathrm{d}t + \sigma\mathrm{d}C_t \tag{5.12}$$

has a solution

$$X_t = \frac{m}{a} + \exp(-at)\left(X_0 - \frac{m}{a}\right) + \sigma\exp(-at)\int_0^t \exp(as)\mathrm{d}C_s \tag{5.13}$$

provided that $a \neq 0$.

5.2   Existence and Uniqueness Theorem

Theorem 5.1 *(Chen and Liu [17], Existence and Uniqueness Theorem) The uncertain differential equation*

$$\mathrm{d}X_t = f(t, X_t)\mathrm{d}t + g(t, X_t)\mathrm{d}C_t \tag{5.14}$$

has a unique solution if the coefficients $f(x, t)$ and $g(x, t)$ satisfy the Lipschitz condition

$$|f(x, t) - f(y, t)| + |g(x, t) - g(y, t)| \leq L|x - y|, \quad \forall x, y \in \Re, t \geq 0 \tag{5.15}$$

and linear growth condition

$$|f(x, t)| + |g(x, t)| \leq L(1 + |x|), \quad \forall x \in \Re, t \geq 0 \tag{5.16}$$

for some constant L. Moreover, the solution is sample-continuous.

Proof: We first prove the existence of solution by a successive approximation method. Define $X_t^{(0)} = X_0$, and

$$X_t^{(n)} = X_0 + \int_0^t f\left(X_s^{(n-1)}, s\right) \mathrm{d}s + \int_0^t g\left(X_s^{(n-1)}, s\right) \mathrm{d}C_s$$

for $n = 1, 2, \cdots$ and write

$$D_t^{(n)}(\gamma) = \max_{0 \le s \le t} \left| X_s^{(n+1)}(\gamma) - X_s^{(n)}(\gamma) \right|$$

for each $\gamma \in \Gamma$. It follows from the Lipschitz condition and linear growth condition that

$$\begin{aligned}
D_t^{(0)}(\gamma) &= \max_{0 \le s \le t} \left| \int_0^s f(X_0, v)\mathrm{d}v + \int_0^s g(X_0, v)\mathrm{d}C_v(\gamma) \right| \\
&\le \int_0^t |f(X_0, v)| \, \mathrm{d}v + K_\gamma \int_0^t |g(X_0, v)| \, \mathrm{d}v \\
&\le (1 + |X_0|)L(1 + K_\gamma)t
\end{aligned}$$

where K_γ is the Lipschitz constant to the sample path $C_t(\gamma)$. In fact, by using the induction method, we may verify

$$D_t^{(n)}(\gamma) \le (1 + |X_0|)\frac{L^{n+1}(1 + K_\gamma)^{n+1}}{(n+1)!} t^{n+1}$$

for each n. This means that, for each sample γ, the paths $X_t^{(k)}(\gamma)$ converges uniformly on any given interval $[0, T]$. Write the limit by $X_t(\gamma)$ that is just a solution of the uncertain differential equation because

$$X_t = X_0 + \int_0^t f(X_s, s)\mathrm{d}s + \int_0^t g(X_s, s)\mathrm{d}s.$$

Next we prove that the solution is unique. Assume that both X_t and X_t^* are solutions of the uncertain differential equation. Then for each $\gamma \in \Gamma$, it follows from the Lipschitz condition and linear growth condition that

$$|X_t(\gamma) - X_t^*(\gamma)| \le L(1 + K_\gamma) \int_0^t |X_v(\gamma) - X_v^*(\gamma)|\mathrm{d}v.$$

By using Gronwall inequality, we obtain

$$|X_t(\gamma) - X_t^*(\gamma)| \le 0 \cdot \exp(L(1 + K_\gamma)t) = 0.$$

Hence $X_t = X_t^*$. The uniqueness is proved. Finally, let us prove the sample-continuity of X_t. The Lipschitz condition and linear growth condition may

produce

$$|X_t(\gamma) - X_s(\gamma)| = \left| \int_s^t f(X_v(\gamma), v)\mathrm{d}v + \int_s^t g(X_v(\gamma), v)\mathrm{d}C_v(\gamma) \right|$$
$$\leq (1 + K_\gamma)(1 + |X_0|)\exp(L(1 + K_\gamma)t)(t - s)$$
$$\to 0 \text{ as } s \to t.$$

Thus X_t is sample-continuous and the theorem is proved.

5.3 Stability

Definition 5.2 *(Liu [130]) An uncertain differential equation is said to be stable if for any given $\kappa > 0$ and $\varepsilon > 0$, there exists $\delta > 0$ such that for any solutions X_t and Y_t, we have*

$$\mathcal{M}\{|X_t - Y_t| > \kappa\} < \varepsilon, \quad \forall t > 0 \tag{5.17}$$

whenever $|X_0 - Y_0| < \delta$.

In other words, an uncertain differential equation is stable if for any given $\kappa > 0$, we have

$$\lim_{|X_0 - Y_0| \to 0} \mathcal{M}\{|X_t - Y_t| > \kappa\} = 0, \quad \forall t > 0. \tag{5.18}$$

Example 5.6: The uncertain differential equation $\mathrm{d}X_t = a\mathrm{d}t + b\mathrm{d}C_t$ is stable since for any given $\kappa > 0$ and $\varepsilon > 0$, we may take $\delta = \kappa$ and have

$$\mathcal{M}\{|X_t - Y_t| > \kappa\} = \mathcal{M}\{|X_0 - Y_0| > \kappa\} = \mathcal{M}\{\emptyset\} = 0 < \varepsilon$$

for any $t > 0$ whenever $|X_0 - Y_0| < \delta$.

Example 5.7: The uncertain differential equation $\mathrm{d}X_t = X_t\mathrm{d}t + b\mathrm{d}C_t$ is unstable since for any given $\kappa > 0$ and any different initial solutions X_0 and Y_0, we have

$$\mathcal{M}\{|X_t - Y_t| > \kappa\} = \mathcal{M}\{\exp(t)|X_0 - Y_0| > \kappa\} = 1$$

provided that t is sufficiently large.

Theorem 5.2 *(Chen [20], Stability Theorem) Suppose u_t and v_t are continuous functions such that*

$$\sup_{s \geq 0} \int_0^s u_t\mathrm{d}t < +\infty, \quad \int_0^{+\infty} |v_t|\mathrm{d}t < +\infty. \tag{5.19}$$

Then the uncertain differential equation

$$\mathrm{d}X_t = u_tX_t\mathrm{d}t + v_tX_t\mathrm{d}C_t \tag{5.20}$$

is stable.

Proof: It has been proved that the unique solution of the uncertain differential equation $\mathrm{d}X_t = u_t X_t \mathrm{d}t + v_t X_t \mathrm{d}C_t$ is

$$X_t = X_0 \exp\left(\int_0^t u_s \mathrm{d}s + \int_0^t v_s \mathrm{d}C_s\right).$$

Thus for any given $\kappa > 0$, we have

$$\mathcal{M}\{|X_t - Y_t| > \kappa\} = \mathcal{M}\left\{|X_0 - Y_0|\exp\left(\int_0^t u_s \mathrm{d}s + \int_0^t v_s \mathrm{d}C_s\right) > \kappa\right\}$$

$$= \mathcal{M}\left\{\int_0^t v_s \mathrm{d}C_s > \ln\frac{\kappa}{|X_0 - Y_0|} - \int_0^t u_s \mathrm{d}s\right\} \to 0$$

as $|X_0 - Y_0| \to 0$ because

$$\int_0^t v_s \mathrm{d}C_s \sim \mathcal{N}\left(0, \int_0^t |v_s|\mathrm{d}s\right)$$

is a normal uncertain variable with expected value 0 and finite variance, and

$$\ln\frac{\kappa}{|X_0 - Y_0|} - \int_0^t u_s \mathrm{d}s \to +\infty.$$

The theorem is proved.

Theorem 5.3 *(Chen [20], Stability Theorem) Suppose u_t and v_t are continuous functions such that*

$$\sup_{s \geq 0}\left|\int_0^s u_t \mathrm{d}t\right| < +\infty, \quad \int_0^{+\infty} |v_t|\mathrm{d}t < +\infty. \tag{5.21}$$

Then the uncertain differential equation

$$\mathrm{d}X_t = u_t X_t \mathrm{d}t + v_t \mathrm{d}C_t \tag{5.22}$$

is stable.

Proof: It has been proved that the unique solution of the uncertain differential equation $\mathrm{d}X_t = u_t X_t \mathrm{d}t + v_t \mathrm{d}C_t$ is

$$X_t = \exp\left(X_0 + \int_0^t u_s \mathrm{d}s + \int_0^t v_s \exp\left(-\int_0^s u_r \mathrm{d}r\right)\mathrm{d}C_s\right).$$

Thus for any given $\kappa > 0$, the uncertain measure $\mathcal{M}\{|X_t - Y_t| > \kappa\}$ is just

$$\mathcal{M}\left\{\int_0^t v_s \exp\left(-\int_0^s u_r \mathrm{d}r\right)\mathrm{d}C_s > \ln\frac{\kappa}{|\exp(X_0) - \exp(Y_0)|} - \int_0^t u_s \mathrm{d}s\right\} \to 0$$

as $|X_0 - Y_0| \to 0$ because

$$\int_0^t v_s \exp\left(-\int_0^s u_r \mathrm{d}r\right) \mathrm{d}C_s \sim \mathcal{N}\left(0, \int_0^t |v_s| \exp\left(-\int_0^s u_r \mathrm{d}r\right) \mathrm{d}s\right)$$

is a normal uncertain variable with expected value 0 and finite variance, and

$$\ln \frac{\kappa}{|\exp(X_0) - \exp(Y_0)|} - \int_0^t u_s \mathrm{d}s \to +\infty.$$

The theorem is proved.

5.4   Uncertain Finance

If we assume that the stock price follows some uncertain differential equation, then we may produce a new topic of uncertain finance. As an example, Liu [130] supposed that the stock price follows geometric canonical process and presented a *stock model* in which the bond price X_t and the stock price Y_t are determined by

$$\begin{cases} \mathrm{d}X_t = rX_t\mathrm{d}t \\ \mathrm{d}Y_t = eY_t\mathrm{d}t + \sigma Y_t\mathrm{d}C_t \end{cases} \qquad (5.23)$$

where r is the riskless interest rate, e is the stock drift, σ is the stock diffusion, and C_t is a canonical process.

European Call Option Price

A European call option gives the holder the right to buy a stock at a specified time for specified price. Assume that the option has strike price K and expiration time s. Then the payoff from such an option is $(Y_s - K)^+$. Considering the time value of money resulted from the bond, the present value of this payoff is $\exp(-rs)(Y_s - K)^+$. Hence the European call option price should be the expected present value of the payoff,

$$f_c = \exp(-rs)E[(Y_s - K)^+]. \qquad (5.24)$$

It is clear that the option price is a decreasing function of interest rate r. That is, the European call option will devaluate if the interest rate is raised; and the European call option will appreciate in value if the interest rate is reduced. In addition, the option price is also a decreasing function of strike price K.

Let us consider the financial market described by the stock model (5.23).
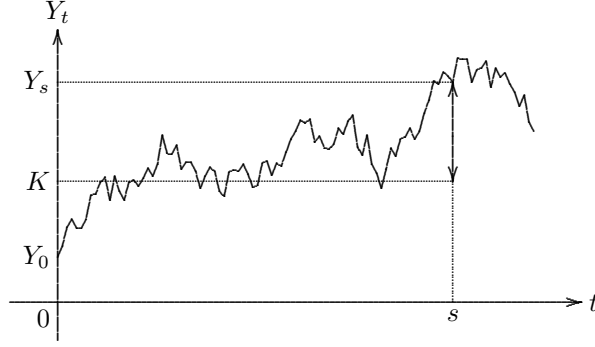
Figure 5.1: Payoff $(Y_s - K)^+$ from European Call Option

The European call option price is

$$
\begin{aligned}
f_c &= \exp(-rs)E[(Y_0 \exp(es + \sigma C_s) - K)^+] \\
&= \exp(-rs) \int_0^{+\infty} \mathcal{M}\{Y_0 \exp(es + \sigma C_s) - K \geq x\} \mathrm{d}x \\
&= \exp(-rs)Y_0 \int_{K/Y_0}^{+\infty} \mathcal{M}\{\exp(es + \sigma C_s) \geq y\} \mathrm{d}y \\
&= \exp(-rs)Y_0 \int_{K/Y_0}^{+\infty} \mathcal{M}\{es + \sigma C_s \geq \ln y\} \mathrm{d}y \\
&= \exp(-rs)Y_0 \int_{K/Y_0}^{+\infty} \left(1 + \exp\left(\frac{\pi(\ln y - es)}{\sqrt{3}\sigma s}\right)\right)^{-1} \mathrm{d}y.
\end{aligned}
$$

Thus Liu [130] derived the following European call option price formula,

$$
f_c = \exp(-rs)Y_0 \int_{K/Y_0}^{+\infty} \left(1 + \exp\left(\frac{\pi(\ln y - es)}{\sqrt{3}\sigma s}\right)\right)^{-1} \mathrm{d}y. \tag{5.25}
$$

European Put Option Price

A European put option gives the holder the right to sell a stock at a specified time for specified price. Assume that the option has strike price K and expiration time s. Then the payoff from such an option is $(K - Y_s)^+$. Considering the time value of money resulted from the bond, the present value of this payoff is $\exp(-rs)(K - Y_s)^+$. Hence the European put option price should be the expected present value of the payoff,

$$
f_p = \exp(-rs)E[(K - Y_s)^+]. \tag{5.26}
$$

It is easy to verify that the option price is a decreasing function of interest rate r, and is an increasing function of strike price K.

Let us consider the financial market described by the stock model (5.23). The European put option price is

$$
\begin{aligned}
f_p &= \exp(-rs)E[(K - Y_0 \exp(es + \sigma C_s))^+] \\
&= \exp(-rs) \int_0^{+\infty} \mathcal{M}\{K - Y_0 \exp(es + \sigma C_s) \geq x\} \mathrm{d}x \\
&= \exp(-rs)Y_0 \int_{K/Y_0}^{+\infty} \mathcal{M}\{\exp(es + \sigma C_s) \leq y\} \mathrm{d}y \\
&= \exp(-rs)Y_0 \int_0^{K/Y_0} \mathcal{M}\{es + \sigma C_s \leq \ln y\} \mathrm{d}y \\
&= \exp(-rs)Y_0 \int_0^{K/Y_0} \left(1 + \exp\left(\frac{\pi(es - \ln y)}{\sqrt{3}\sigma s}\right)\right)^{-1} \mathrm{d}y.
\end{aligned}
$$

Thus Liu [130] derived the following European put option price formula,

$$
f_p = \exp(-rs)Y_0 \int_0^{K/Y_0} \left(1 + \exp\left(\frac{\pi(es - \ln y)}{\sqrt{3}\sigma s}\right)\right)^{-1} \mathrm{d}y. \qquad (5.27)
$$

Multi-factor Stock Model

Now we assume that there are multiple stocks whose prices are determined by multiple canonical processes. For this case, we have a *multi-factor stock model* in which the bond price X_t and the stock prices Y_{it} are determined by

$$
\begin{cases}
\mathrm{d}X_t = rX_t\mathrm{d}t \\[2mm]
\mathrm{d}Y_{it} = e_i Y_{it}\mathrm{d}t + \sum_{j=1}^{n} \sigma_{ij}Y_{it}\mathrm{d}C_{jt}, \ \ i = 1, 2, \cdots, m
\end{cases}
\qquad (5.28)
$$

where r is the riskless interest rate, e_i are the stock drift coefficients, σ_{ij} are the stock diffusion coefficients, C_{it} are independent canonical processes, $i = 1, 2, \cdots, m$, $j = 1, 2, \cdots, n$.

Portfolio Selection

For the stock model (5.28), we have the choice of $m+1$ different investments. At each instant t we may choose a portfolio $(\beta_t, \beta_{1t}, \cdots, \beta_{mt})$ (i.e., the investment fractions meeting $\beta_t + \beta_{1t} + \cdots + \beta_{mt} = 1$). Then the wealth Z_t at time t should follow the uncertain differential equation

$$
\mathrm{d}Z_t = r\beta_t Z_t\mathrm{d}t + \sum_{i=1}^{m} e_i\beta_{it}Z_t\mathrm{d}t + \sum_{i=1}^{m}\sum_{j=1}^{n} \sigma_{ij}\beta_{it}Z_t\mathrm{d}C_{jt}. \qquad (5.29)
$$

Portfolio selection problem is to find an optimal portfolio $(\beta_t, \beta_{1t}, \cdots, \beta_{mt})$ such that the expected wealth $E[Z_s]$ is maximized.

No-Arbitrage

The stock model (5.28) is said to be *no-arbitrage* if there is no portfolio $(\beta_t, \beta_{1t}, \cdots, \beta_{mt})$ such that for some time $s > 0$, we have

$$\mathcal{M}\{\exp(-rs)Z_s \geq Z_0\} = 1 \tag{5.30}$$

and

$$\mathcal{M}\{\exp(-rs)Z_s > Z_0\} > 0 \tag{5.31}$$

where Z_t is determined by (5.29) and represents the wealth at time t. We may prove that the stock model (5.28) is no-arbitrage if and only if its diffusion matrix

$$\begin{pmatrix} \sigma_{11} & \sigma_{12} & \cdots & \sigma_{1n} \\ \sigma_{21} & \sigma_{22} & \cdots & \sigma_{2n} \\ \vdots & \vdots & \ddots & \vdots \\ \sigma_{m1} & \sigma_{m2} & \cdots & \sigma_{mn} \end{pmatrix}$$

has rank m, i.e., the row vectors are linearly independent.

Stock Model with Periodic Dividends

Liu [140] assumed that a dividend of a fraction δ of the stock price is paid at deterministic times $T_1, T_2, \cdots$ and presented a stock model with periodic dividends,

$$\begin{cases} X_t = X_0 \exp(rt) \\ Y_t = Y_0(1-\delta)^{n_t} \exp(et + \sigma C_t) \end{cases} \tag{5.32}$$

where $n_t = \max\{i : T_i \leq t\}$ is the number of dividend payments made by time t.

Stock Model with Mean-Reverting Process

Peng [184] assumed that the stock price follows a type of mean-reverting uncertain process and proposed the following stock model,

$$\begin{cases} \mathrm{d}X_t = rX_t\mathrm{d}t \\ \mathrm{d}Y_t = (m - \alpha Y_t)\mathrm{d}t + \sigma \mathrm{d}C_t \end{cases} \tag{5.33}$$

where r, m, α, σ are given constants.

Currency Models

Liu [139] assumed that the exchange rate follows a geometric canonical process and proposed a currency model with uncertain exchange rate,

$$\begin{cases} \mathrm{d}X_t = eX_t\mathrm{d}t + \sigma X_t\mathrm{d}C_t & \text{(Exchange rate)} \\ \mathrm{d}Y_t = uY_t\mathrm{d}t & \text{(Yuan Bond)} \\ \mathrm{d}Z_t = vZ_t\mathrm{d}t & \text{(Dollar Bond)} \end{cases} \tag{5.34}$$

where e, σ, u, v are constants. If the exchange rate follows a mean-reverting uncertain process, then the currency model with uncertain exchange rate is

$$
\begin{cases}
\mathrm{d}X_t = (m - \alpha X_t)\mathrm{d}t + \sigma \mathrm{d}C_t & \text{(Exchange rate)} \\
\mathrm{d}Y_t = u Y_t \mathrm{d}t & \text{(Yuan Bond)} \\
\mathrm{d}Z_t = v Z_t \mathrm{d}t & \text{(Dollar Bond)}
\end{cases}
\tag{5.35}
$$

where m, α, σ, u, v are constants.

5.5 Uncertain Filtering

One outstanding contribution to filtering problem is Kalman-Bucy filter when the noise is assumed to be Brownian motion. If we suppose that the noise is a canonical process instead of Brownian motion, how do we filter the noise away from the observation?

Suppose an uncertain system X_t is described by an uncertain differential equation

$$
\mathrm{d}X_t = f(t)X_t\mathrm{d}t + u(t)\mathrm{d}C_t
\tag{5.36}
$$

where C_t is a canonical process. We also have an observation Y_t with

$$
\mathrm{d}Y_t = g(t)X_t\mathrm{d}t + v(t)\mathrm{d}D_t
\tag{5.37}
$$

where D_t is another canonical process that is independent of C_t. The *filtering problem* is to find the best estimate of X_t based on the observation Y_t.

5.6 Uncertain Optimal Control

An *uncertain optimal control system* is assumed to follow an uncertain differential equation

$$
\mathrm{d}X_t = f(t, X_t, Y_t)\mathrm{d}t + g(t, X_t, Y_t)\mathrm{d}C_t
\tag{5.38}
$$

where X_t is the state, Y_t is a control, and C_t is a canonical process. Assume that R is the return function and T is the function of terminal reward. If we want to maximize the expected return on $[0, s]$ by using an optimal control, then we have the following uncertain optimal control model,

$$
\begin{cases}
\max\limits_{Y_t} E\left[\displaystyle\int_0^s R(t, X_t, Y_t)\mathrm{d}t + T(s, X_s)\right] \\
\text{subject to:} \\
\quad \mathrm{d}X_t = f(t, X_t, Y_t)\mathrm{d}t + g(t, X_t, Y_t)\mathrm{d}C_t.
\end{cases}
\tag{5.39}
$$

If we say Hamilton-Jacobi-Bellman equation provides a necessary condition for extremum of stochastic optimal control model, then Zhu's equation of optimality [250] provides a necessary condition for extremum of uncertain optimal control model.

5.7   Uncertain Differential Equation with Jumps

In many cases the stock price is not continuous because of economic crisis or war. In order to incorporate those into stock model, we should develop an uncertain calculus with jump process. For many applications, a renewal process N_t is sufficient. The uncertain integral of uncertain process X_t with respect to N_t is

$$\int_a^b X_t \mathrm{d}N_t = \lim_{\Delta \to 0} \sum_{i=1}^k X_{t_i} \cdot (N_{t_{i+1}} - N_{t_i}) = \sum_{a \le t \le b} X_t \cdot (N_t - N_{t-}). \quad (5.40)$$

Definition 5.3 *Suppose C_t is a canonical process, N_t is a renewal process, and f, g, h are some given functions. Then*

$$\mathrm{d}X_t = f(t, X_t)\mathrm{d}t + g(t, X_t)\mathrm{d}C_t + h(t, X_t)\mathrm{d}N_t \quad (5.41)$$

is called an uncertain differential equation with jumps. A solution is an uncertain process X_t that satisfies (5.41) identically in t.

Example 5.8: Let C_t be a canonical process and N_t a renewal process. Then the uncertain differential equation with jumps

$$\mathrm{d}X_t = a\mathrm{d}t + b\mathrm{d}C_t + c\mathrm{d}N_t$$

has a solution $X_t = at + bC_t + cN_t$ which is just a jump process.

Example 5.9: Let C_t be a canonical process and N_t a renewal process. Then the uncertain differential equation with jumps

$$\mathrm{d}X_t = aX_t\mathrm{d}t + bX_t\mathrm{d}C_t + cX_t\mathrm{d}N_t$$

has a solution $X_t = \exp\left(at + bC_t + cN_t\right)$ which may be employed to model stock price with jumps.

Chapter 6

Uncertain Logic

Uncertain logic was designed by Li and Liu [99] in 2009 as a generalization of mathematical logic for dealing with uncertain knowledge via uncertainty theory. A key point in uncertain logic is that the truth value of an uncertain proposition is defined as the uncertain measure that the proposition is true. Another important contribution is the truth value theorem by Chen and Ralescu [18] that provides a numerical method for calculating the truth value of uncertain formulas.

One advantage of uncertain logic is the well consistency with classical logic. For example, uncertain logic is consistent with the law of excluded middle and law of contradiction.

6.1   Uncertain Proposition

Definition 6.1 *An uncertain proposition is a statement whose truth value is quantified by an uncertain measure.*

That is, if we use ξ to express the uncertain proposition and use c to express its truth value in uncertain measure, then the uncertain proposition ξ is essentially an uncertain variable

$$\xi = \begin{cases} 1 \text{ with uncertain measure } c \\ 0 \text{ with uncertain measure } 1 - c \end{cases} \tag{6.1}$$

where $\xi = 1$ means ξ is true and $\xi = 0$ means ξ is false.

Example 6.1: "Tom is tall with truth value 0.7" is an uncertain proposition, where "Tom is tall" is a statement, and its truth value is 0.7 in uncertain measure.

Example 6.2: "Beijing is a big city with truth value 0.9" is an uncertain

proposition, where "Beijing is a big city" is a statement, and its truth value is 0.9 in uncertain measure.

Definition 6.2 *Uncertain propositions are called independent if they are independent uncertain variables.*

Example 6.3: If ξ and η are independent uncertain propositions, then for any x and y taking values either 0 or 1, we have

$$\mathcal{M}\{(\xi = x) \cap (\eta = y)\} = \mathcal{M}\{\xi = x\} \wedge \mathcal{M}\{\eta = y\},$$
$$\mathcal{M}\{(\xi = x) \cup (\eta = y)\} = \mathcal{M}\{\xi = x\} \vee \mathcal{M}\{\eta = y\}.$$

6.2 Connective Symbols

In addition to the proposition symbols ξ and η, we also need the negation symbol $\neg$, conjunction symbol $\wedge$, disjunction symbol $\vee$, conditional symbol $\rightarrow$, and biconditional symbol $\leftrightarrow$. Note that

$$\neg \xi \text{ means "not } \xi\text{"}; \tag{6.2}$$

$$\xi \vee \eta \text{ means "}\xi \text{ or } \eta\text{"}; \tag{6.3}$$

$$\xi \wedge \eta = \neg(\neg\xi \vee \neg\eta) \text{ means "}\xi \text{ and } \eta\text{"}; \tag{6.4}$$

$$\xi \rightarrow \eta = (\neg\xi) \vee \eta \text{ means "if } \xi \text{ then } \eta\text{"}; \tag{6.5}$$

$$\xi \leftrightarrow \eta = (\xi \rightarrow \eta) \wedge (\eta \rightarrow \xi) \text{ means "}\xi \text{ if and only if } \eta\text{"}. \tag{6.6}$$

6.3 Uncertain Formula

An *uncertain formula* is a finite sequence of uncertain propositions and connective symbols that must make sense. For example, let ξ, η, τ be uncertain propositions. Then

$$X = \neg\xi, \quad X = \xi \wedge \eta, \quad X = (\xi \vee \eta) \rightarrow \tau$$

are all uncertain formulas. However, $\neg \vee \xi$, $\xi \rightarrow \vee$ and $\xi\eta \rightarrow \tau$ are not formulas. Note that an uncertain formula X is essentially an uncertain variable taking values 0 or 1. If $X = 1$, then X is true; if $X = 0$, then X is false.

Definition 6.3 *Uncertain formulas are called independent if they are independent uncertain variables.*

6.4   Truth Function

Assume X is a formula containing propositions $\xi_1, \xi_2, \cdots, \xi_n$. It is well-known that there is a function $f : \{0,1\}^n \to \{0,1\}$ such that $X = 1$ if and only if $f(\xi_1, \xi_2, \cdots, \xi_n) = 1$. Such a Boolean function f is called the *truth function* of X.

Example 6.4: The truth function of formula $\xi_1 \vee \xi_2$ (ξ_1 or ξ_2) is

$$f(1,1) = 1, \quad f(1,0) = 1, \quad f(0,1) = 1, \quad f(0,0) = 0.$$

Example 6.5: The truth function of formula $\xi_1 \wedge \xi_2$ (ξ_1 and ξ_2) is

$$f(1,1) = 1, \quad f(1,0) = 0, \quad f(0,1) = 0, \quad f(0,0) = 0.$$

Example 6.6: The truth function of formula $\xi_1 \to \xi_2$ (if ξ_1 then ξ_2) is

$$f(1,1) = 1, \quad f(1,0) = 0, \quad f(0,1) = 1, \quad f(0,0) = 1.$$

Example 6.7: The truth function of $\xi_1 \leftrightarrow \xi_2$ (ξ_1 if and only if ξ_2) is

$$f(1,1) = 1, \quad f(1,0) = 0, \quad f(0,1) = 0, \quad f(0,0) = 1.$$

Example 6.8: The truth function of $\xi_1 \vee \xi_2 \to \xi_3$ is given by

$$f(1,1,1) = 1, \quad f(1,0,1) = 1, \quad f(0,1,1) = 1, \quad f(0,0,1) = 1,$$
$$f(1,1,0) = 0, \quad f(1,0,0) = 0, \quad f(0,1,0) = 0, \quad f(0,0,0) = 1.$$

6.5   Truth Value

Truth value is a key concept in uncertain logic, and is defined as the uncertain measure that the uncertain formula is true.

Definition 6.4 *(Li and Liu [99]) Let X be an uncertain formula. Then the truth value of X is defined as the uncertain measure that the uncertain formula X is true, i.e.,*

$$T(X) = \mathcal{M}\{X = 1\}. \tag{6.7}$$

The truth value is nothing but an uncertain measure. The truth value $T(X) = 1$ means X is certainly true, $T(X) = 0$ means X is certainly false, and $T(X) = 0.5$ means X is totally uncertain. The higher the truth value is, the more true the uncertain formula is.

Example 6.9: Let ξ and η be two independent uncertain propositions with truth values a and b, respectively. Then

$$T(\xi) = \mathcal{M}\{\xi = 1\} = a, \tag{6.8}$$

$$T(\neg\xi) = \mathcal{M}\{\xi = 0\} = 1 - a, \tag{6.9}$$

$$T(\xi \vee \eta) = \mathcal{M}\{\xi \vee \eta = 1\} = \mathcal{M}\{(\xi = 1) \cup (\eta = 1)\} = a \vee b, \tag{6.10}$$

$$T(\xi \wedge \eta) = \mathcal{M}\{\xi \wedge \eta = 1\} = \mathcal{M}\{(\xi = 1) \cap (\eta = 1)\} = a \wedge b, \tag{6.11}$$

$$T(\xi \to \eta) = T(\neg\xi \vee \eta) = (1 - a) \vee b. \tag{6.12}$$

6.6   Laws

Theorem 6.1 *(Law of Excluded Middle) For any uncertain formula X, we have*

$$T(X \vee \neg X) = 1. \tag{6.13}$$

Proof: It follows from the definition of truth value and property of uncertain measure that

$$T(X \vee \neg X) = \mathcal{M}\{X \vee \neg X = 1\} = \mathcal{M}\{(X = 1) \cup (X = 0)\} = 1.$$

The theorem is proved.

Theorem 6.2 *(Law of Contradiction) For any uncertain formula X, we have*

$$T(X \wedge \neg X) = 0. \tag{6.14}$$

Proof: It follows from the definition of truth value and property of uncertain measure that

$$T(X \wedge \neg X) = \mathcal{M}\{X \wedge \neg X = 1\} = \mathcal{M}\{(X = 1) \cap (X = 0)\} = \mathcal{M}\{\emptyset\} = 0.$$

The theorem is proved.

Theorem 6.3 *(Law of Truth Conservation) For any uncertain formula X, we have*

$$T(X) + T(\neg X) = 1. \tag{6.15}$$

Proof: It follows from the self-duality of uncertain measure that

$$T(\neg X) = \mathcal{M}\{\neg X = 1\} = \mathcal{M}\{X = 0\} = 1 - \mathcal{M}\{X = 1\} = 1 - T(X).$$

The theorem is proved.

Theorem 6.4 *(De Morgan's Law) For any uncertain formulas X and Y, we have*

$$T(\neg(X \wedge Y)) = T((\neg X) \vee (\neg Y)), \tag{6.16}$$

$$T(\neg(X \vee Y)) = T((\neg X) \wedge (\neg Y)). \tag{6.17}$$

Proof: It follows from the basic properties of uncertain measure that

$$T(\neg(X \wedge Y)) = \mathcal{M}\{X \wedge Y = 0\} = \mathcal{M}\{(X = 0) \cup (Y = 0)\}$$
$$= \mathcal{M}\{(\neg X) \vee (\neg Y) = 1\} = T((\neg X) \vee (\neg Y))$$

which proves the first equality. A similar way may verify the second equality. The theorem is proved.

Theorem 6.5 *(Law of Contraposition) For any uncertain formulas X and Y, we have*

$$T(X \rightarrow Y) = T(\neg Y \rightarrow \neg X). \qquad (6.18)$$

Proof: It follows from the basic properties of uncertain measure that

$$T(X \rightarrow Y) = \mathcal{M}\{(\neg X) \vee Y = 1\} = \mathcal{M}\{(X = 0) \cup (Y = 1)\}$$
$$= \mathcal{M}\{Y \vee (\neg X) = 1\} = T(\neg Y \rightarrow \neg X).$$

The theorem is proved.

Theorem 6.6 *(Monotonicity and Subadditivity) For any uncertain formulas X and Y, we have*

$$T(X) \vee T(Y) \leq T(X \vee Y) \leq T(X) + T(Y). \qquad (6.19)$$

Proof: It follows from the monotonicity of uncertain measure that

$$T(X \vee Y) = \mathcal{M}\{X \vee Y = 1\} = \mathcal{M}\{(X = 1) \cup (Y = 1)\}$$
$$\geq \mathcal{M}\{X = 1\} \vee \mathcal{M}\{Y = 1\} = T(X) \vee T(Y).$$

It follows from the subadditivity of uncertain measure that

$$T(X \vee Y) = \mathcal{M}\{X \vee Y = 1\} = \mathcal{M}\{(X = 1) \cup (Y = 1)\}$$
$$\leq \mathcal{M}\{X = 1\} + \mathcal{M}\{Y = 1\} = T(X) + T(Y).$$

The theorem is verified.

Theorem 6.7 *For any uncertain formulas X and Y, we have*

$$T(X) + T(Y) - 1 \leq T(X \wedge Y) \leq T(X) \wedge T(Y). \qquad (6.20)$$

Proof: It follows from the monotonicity of truth value that

$$T(X \wedge Y) = 1 - T(\neg X \vee \neg Y) \leq 1 - T(\neg X) \vee T(\neg Y)$$
$$= (1 - T(\neg X)) \wedge (1 - T(\neg Y)) = T(X) \wedge T(Y).$$

It follows from the subadditivity of truth value that

$$T(X \wedge Y) = 1 - T(\neg X \vee \neg Y) \geq 1 - (T(\neg X) + T(\neg Y))$$
$$= 1 - (1 - T(X)) - (1 - T(Y)) = T(X) + T(Y) - 1.$$

The theorem is proved.

Theorem 6.8 *For any uncertain formula X, we have*

$$T(X \to X) = 1, \tag{6.21}$$

$$T(X \leftrightarrow X) = 1, \tag{6.22}$$

$$T(X \to \neg X) = 1 - T(X). \tag{6.23}$$

Proof: It follows from the law of excluded middle and law of truth conservation that

$$T(X \to X) = T(\neg X \vee X) = 1,$$

$$T(X \leftrightarrow X) = T((X \to X) \wedge (X \to X)) = T(X \to X) = 1,$$

$$T(X \to \neg X) = T(\neg X \vee \neg X) = T(\neg X) = 1 - T(X).$$

The theorem is proved.

Independence Case

Theorem 6.9 *If two uncertain formulas X and Y are independent, then we have*

$$T(X \vee Y) = T(X) \vee T(Y), \quad T(X \wedge Y) = T(X) \wedge T(Y). \tag{6.24}$$

Proof: Since X and Y are independent uncertain formulas, they are independent uncertain variables. Hence

$$T(X \vee Y) = \mathcal{M}\{X \vee Y = 1\} = \mathcal{M}\{X = 1\} \vee \mathcal{M}\{Y = 1\} = T(X) \vee T(Y),$$
$$T(X \wedge Y) = \mathcal{M}\{X \wedge Y = 1\} = \mathcal{M}\{X = 1\} \wedge \mathcal{M}\{Y = 1\} = T(X) \wedge T(Y).$$

The theorem is proved.

Theorem 6.10 *If two uncertain formulas X and Y are independent, then we have*

$$T(X \to Y) = (1 - T(X)) \vee T(Y). \tag{6.25}$$

Proof: Since X and Y are independent, the formulas $\neg X$ and Y are also independent. It follows that

$$T(X \to Y) = T(\neg X \vee Y) = T(\neg X) \vee T(Y) = (1 - T(X)) \vee T(Y)$$

which proves the theorem.

6.7 Truth Value Theorem

Theorem 6.11 *(Chen and Ralescu [18], Truth Value Theorem) Assume that $\xi_1, \xi_2, \cdots, \xi_n$ are independent uncertain propositions with truth values $a_1, a_2, \cdots, a_n$, respectively. If X is an uncertain formula containing $\xi_1, \xi_2, \cdots, \xi_n$ with truth function f, then the truth value of X is*

$$
T(X) = \begin{cases} \sup\limits_{f(x_1,x_2,\cdots,x_n)=1} \min\limits_{1 \le i \le n} \nu_i(x_i), \\ \qquad if \sup\limits_{f(x_1,x_2,\cdots,x_n)=1} \min\limits_{1 \le i \le n} \nu_i(x_i) < 0.5 \\ 1 - \sup\limits_{f(x_1,x_2,\cdots,x_n)=0} \min\limits_{1 \le i \le n} \nu_i(x_i), \\ \qquad if \sup\limits_{f(x_1,x_2,\cdots,x_n)=1} \min\limits_{1 \le i \le n} \nu_i(x_i) \ge 0.5 \end{cases} \tag{6.26}
$$

where x_i take values either 0 or 1, and v_i are defined by

$$
\nu_i(x_i) = \begin{cases} a_i, & if \ x_i = 1 \\ 1 - a_i, & if \ x_i = 0 \end{cases} \tag{6.27}
$$

for $i = 1, 2, \cdots, n$, respectively.

Proof: Since $X = 1$ if and only if $f(\xi_1, \xi_2, \cdots, \xi_n) = 1$, we immediately have

$$
T(X) = \mathcal{M}\{X = 1\} = \mathcal{M}\{f(\xi_1, \xi_2, \cdots, \xi_n) = 1\}.
$$

It follows from the operational law and independence of uncertain propositions that

$$
T(X) = \begin{cases} \sup\limits_{f(B_1,B_2,\cdots,B_n)=1} \min\limits_{1 \le i \le n} \mathcal{M}\{\xi_i \in B_i\}, \\ \qquad if \sup\limits_{f(B_1,B_2,\cdots,B_n)=1} \min\limits_{1 \le i \le n} \mathcal{M}\{\xi_i \in B_i\} > 0.5 \\ 1 - \sup\limits_{f(B_1,B_2,\cdots,B_n)=0} \min\limits_{1 \le i \le n} \mathcal{M}\{\xi_i \in B_i\}, \\ \qquad if \sup\limits_{f(B_1,B_2,\cdots,B_n)=0} \min\limits_{1 \le i \le n} \mathcal{M}\{\xi_i \in B_i\} > 0.5 \\ 0.5, \quad otherwise \end{cases}
$$

where $B_1, B_2, \cdots, B_n$ are subsets of $\{0, 1\}$, and $f(B_1, B_2, \cdots, B_n) = 1$ means $f(x_1, x_2, \cdots, x_n) = 1$ for any $x_1 \in B_1, x_2 \in B_2, \cdots, x_n \in B_n$. Please also note that

$$
\nu_i(1) = \mathcal{M}\{\xi_i = 1\}, \quad \nu_i(0) = \mathcal{M}\{\xi_i = 0\}
$$

for $i = 1, 2, \cdots, n$. The argument breaks down into three cases. Case 1: Assume

$$
\sup\limits_{f(x_1,x_2,\cdots,x_n)=1} \min\limits_{1 \le i \le n} \nu_i(x_i) < 0.5.
$$

Then we have

$$\sup_{f(B_1,B_2,\cdots,B_n)=0} \min_{1\leq i\leq n} \mathcal{M}\{\xi_i \in B_i\} > 0.5.$$

Thus

$$T(X) = 1 - \sup_{f(B_1,B_2,\cdots,B_n)=0} \min_{1\leq i\leq n} \mathcal{M}\{\xi_i \in B_i\} = \sup_{f(x_1,x_2,\cdots,x_n)=1} \min_{1\leq i\leq n} \nu_i(x_i).$$

Case 2: Assume

$$\sup_{f(x_1,x_2,\cdots,x_n)=1} \min_{1\leq i\leq n} \nu_i(x_i) > 0.5.$$

Then we have

$$\sup_{f(B_1,B_2,\cdots,B_n)=1} \min_{1\leq i\leq n} \mathcal{M}\{\xi_i \in B_i\} \geq \sup_{f(x_1,x_2,\cdots,x_n)=1} \min_{1\leq i\leq n} \nu_i(x_i) > 0.5.$$

Thus

$$T(X) = \sup_{f(B_1,B_2,\cdots,B_n)=1} \min_{1\leq i\leq n} \mathcal{M}\{\xi_i \in B_i\} = 1 - \sup_{f(x_1,x_2,\cdots,x_n)=0} \min_{1\leq i\leq n} \nu_i(x_i).$$

Case 3: Assume

$$\sup_{f(x_1,x_2,\cdots,x_n)=1} \min_{1\leq i\leq n} \nu_i(x_i) = 0.5.$$

Then we have

$$\sup_{f(B_1,B_2,\cdots,B_n)=1} \min_{1\leq i\leq n} \mathcal{M}\{\xi_i \in B_i\} \geq 0.5$$

and

$$\sup_{f(B_1,B_2,\cdots,B_n)=0} \min_{1\leq i\leq n} \mathcal{M}\{\xi_i \in B_i\} \geq 0.5.$$

Thus

$$T(X) = 0.5 = 1 - \sup_{f(x_1,x_2,\cdots,x_n)=0} \min_{1\leq i\leq n} \nu_i(x_i).$$

The theorem is proved.

Example 6.10: Let $\xi_1, \xi_2, \cdots, \xi_n$ be independent uncertain propositions with truth values $a_1, a_2, \cdots, a_n$, respectively. For any integer k with $1 \leq k \leq n$,

$$X = \text{``at least } k \text{ propositions of } \xi_1, \xi_2, \cdots, \xi_n \text{ are true''} \qquad (6.28)$$

is an uncertain formula. It is obvious that the truth function is

$$f(x_1, x_2, \cdots, x_n) = \begin{cases} 1, & \text{if } x_1 + x_2 + \cdots + x_n \geq k \\ 0, & \text{if } x_1 + x_2 + \cdots + x_n < k. \end{cases}$$

The truth value theorem may produce the truth value $T(X)$. Without loss of generality, we assume $a_1 \geq a_2 \geq \cdots \geq a_n$. When $a_k < 0.5$, we have

$$\sup_{f(x_1,x_2,\cdots,x_n)=1} \min_{1 \leq i \leq n} \nu_i(x_i) = a_k < 0.5.$$

It follows from the truth value theorem that

$$T(X) = \sup_{f(x_1,x_2,\cdots,x_n)=1} \min_{1 \leq i \leq n} \nu_i(x_i) = a_k.$$

When $a_k \geq 0.5$, we have

$$\sup_{f(x_1,x_2,\cdots,x_n)=1} \min_{1 \leq i \leq n} \nu_i(x_i) = a_k \wedge \min_{k+1 \leq i \leq n} (a_i \vee (1 - a_i)) \geq 0.5.$$

It follows from the truth value theorem that

$$T(X) = 1 - \sup_{f(x_1,x_2,\cdots,x_n)=0} \min_{1 \leq i \leq n} \nu_i(x_i) = 1 - (1 - a_k) = a_k.$$

Thus we always have $T(X) = a_k$. That is,

$$T(X) = \text{the } k\text{th largest value of } a_1, a_2, \cdots, a_n. \tag{6.29}$$

When $k = 1$, it is clear that $T(X)$ is just the largest value, i.e.,

$$T(\xi_1 \vee \xi_2 \vee \cdots \vee \xi_n) = a_1 \vee a_2 \vee \cdots \vee a_n. \tag{6.30}$$

When $k = n$, it is clear that $T(X)$ is just the smallest value, i.e.,

$$T(\xi_1 \wedge \xi_2 \wedge \cdots \wedge \xi_n) = a_1 \wedge a_2 \wedge \cdots \wedge a_n. \tag{6.31}$$

Example 6.11: Let ξ_1 and ξ_2 be independent uncertain propositions with truth values a_1 and a_2, respectively. Then

$$X = \xi_1 \leftrightarrow \xi_2 \tag{6.32}$$

is an uncertain formula. It is clear that the truth function is

$$f(1,1) = 1, \quad f(1,0) = 0, \quad f(0,1) = 0, \quad f(0,0) = 1.$$

Let us calculate its truth value $T(X)$ by the truth value theorem. When $a_1 \geq 0.5$ and $a_2 \geq 0.5$, we have

$$\sup_{f(x_1,x_2)=1} \min_{1 \leq i \leq 2} \nu_i(x_i) = \max\{a_1 \wedge a_2, (1 - a_1) \wedge (1 - a_2)\} = a_1 \wedge a_2 \geq 0.5.$$

It follows from the truth value theorem that

$$T(X) = 1 - \sup_{f(x_1,x_2)=0} \min_{1 \leq i \leq 2} \nu_i(x_i) = 1 - \max\{(1-a_1) \wedge a_2, a_1 \wedge (1-a_2)\} = a_1 \wedge a_2.$$

When $a_1 \geq 0.5$ and $a_2 < 0.5$, we have

$$\sup_{f(x_1,x_2)=1} \min_{1 \leq i \leq 2} \nu_i(x_i) = \max\{a_1 \wedge a_2, (1-a_1) \wedge (1-a_2)\} = (1-a_1) \vee a_2 \leq 0.5.$$

It follows from the truth value theorem that

$$T(X) = \sup_{f(x_1,x_2)=1} \min_{1 \leq i \leq 2} \nu_i(x_i) = (1-a_1) \vee a_2.$$

When $a_1 < 0.5$ and $a_2 \geq 0.5$, we have

$$\sup_{f(x_1,x_2)=1} \min_{1 \leq i \leq 2} \nu_i(x_i) = \max\{a_1 \wedge a_2, (1-a_1) \wedge (1-a_2)\} = a_1 \vee (1-a_2) \leq 0.5.$$

It follows from the truth value theorem that

$$T(X) = \sup_{f(x_1,x_2)=1} \min_{1 \leq i \leq 2} \nu_i(x_i) = a_1 \vee (1-a_2).$$

When $a_1 < 0.5$ and $a_2 < 0.5$, we have

$$\sup_{f(x_1,x_2)=1} \min_{1 \leq i \leq 2} \nu_i(x_i) = \max\{a_1 \wedge a_2, (1-a_1) \wedge (1-a_2)\} = (1-a_1) \wedge (1-a_2) > 0.5.$$

It follows from the truth value theorem that

$$T(X) = 1 - \sup_{f(x_1,x_2)=0} \min_{1 \leq i \leq 2} \nu_i(x_i) = 1 - a_1 \vee a_2 = (1-a_1) \wedge (1-a_2).$$

Thus we have

$$T(X) = \begin{cases} a_1 \wedge a_2, & \text{if } a_1 \geq 0.5 \text{ and } a_2 \geq 0.5 \\ (1-a_1) \vee a_2, & \text{if } a_1 \geq 0.5 \text{ and } a_2 < 0.5 \\ a_1 \vee (1-a_2), & \text{if } a_1 < 0.5 \text{ and } a_2 \geq 0.5 \\ (1-a_1) \wedge (1-a_2), & \text{if } a_1 < 0.5 \text{ and } a_2 < 0.5. \end{cases} \tag{6.33}$$

Example 6.12: Let ξ_1 and ξ_2 be independent uncertain propositions with truth values a_1 and a_2, respectively. Then

$$X = \text{``}\xi_1 \text{ or } \xi_2 \text{ and not both''} \tag{6.34}$$

is an uncertain formula. It is obvious that the truth function is

$$f(1,1) = 0, \quad f(1,0) = 1, \quad f(0,1) = 1, \quad f(0,0) = 0.$$

Let us employ the truth value theorem to calculate the truth value $T(X)$. When $a_1 \geq 0.5$ and $a_2 \geq 0.5$, we have

$$\sup_{f(x_1,x_2)=1} \min_{1 \leq i \leq 2} \nu_i(x_i) = \max\{a_1 \wedge (1-a_2), (1-a_1) \wedge a_2\} = (1-a_1) \vee (1-a_2) \leq 0.5.$$

It follows from the truth value theorem that

$$T(X) = \sup_{f(x_1,x_2)=1} \min_{1 \le i \le 2} \nu_i(x_i) = (1 - a_1) \vee (1 - a_2).$$

When $a_1 \ge 0.5$ and $a_2 < 0.5$, we have

$$\sup_{f(x_1,x_2)=1} \min_{1 \le i \le 2} \nu_i(x_i) = \max\{a_1 \wedge (1-a_2), (1-a_1) \wedge a_2\} = a_1 \wedge (1-a_2) \ge 0.5.$$

It follows from the truth value theorem that

$$T(X) = 1 - \sup_{f(x_1,x_2)=0} \min_{1 \le i \le 2} \nu_i(x_i) = 1 - (1-a_1) \vee a_2 = a_1 \wedge (1-a_2).$$

When $a_1 < 0.5$ and $a_2 \ge 0.5$, we have

$$\sup_{f(x_1,x_2)=1} \min_{1 \le i \le 2} \nu_i(x_i) = \max\{a_1 \wedge (1-a_2), (1-a_1) \wedge a_2\} = (1-a_1) \wedge a_2 \ge 0.5.$$

It follows from the truth value theorem that

$$T(X) = 1 - \sup_{f(x_1,x_2)=0} \min_{1 \le i \le 2} \nu_i(x_i) = 1 - a_1 \vee (1-a_2) = (1-a_1) \wedge a_2.$$

When $a_1 < 0.5$ and $a_2 < 0.5$, we have

$$\sup_{f(x_1,x_2)=1} \min_{1 \le i \le 2} \nu_i(x_i) = \max\{a_1 \wedge (1-a_2), (1-a_1) \wedge a_2\} = a_1 \vee a_2 < 0.5.$$

It follows from the truth value theorem that

$$T(X) = \sup_{f(x_1,x_2)=1} \min_{1 \le i \le 2} \nu_i(x_i) = a_1 \vee a_2.$$

Thus we have

$$T(X) = \begin{cases} (1-a_1) \vee (1-a_2), & \text{if } a_1 \ge 0.5 \text{ and } a_2 \ge 0.5 \\ a_1 \wedge (1-a_2), & \text{if } a_1 \ge 0.5 \text{ and } a_2 < 0.5 \\ (1-a_1) \wedge a_2, & \text{if } a_1 < 0.5 \text{ and } a_2 \ge 0.5 \\ a_1 \vee a_2, & \text{if } a_1 < 0.5 \text{ and } a_2 < 0.5. \end{cases} \tag{6.35}$$

6.8   Truth Value Solver

Truth Value Solver is a software for computing the truth values of uncertain formula based on the truth value theorem. This software may be downloaded from http://orsc.edu.cn/liu/resources.htm. Now let us perform it via some numerical examples.

Example 6.13: Assume that $\xi_1, \xi_2, \xi_3, \xi_4, \xi_5$ are independent uncertain propositions with truth values $0.1, 0.3, 0.5, 0.7, 0.9$, respectively. Let

$$X = (\xi_1 \wedge \xi_2) \vee (\xi_2 \wedge \xi_3) \vee (\xi_3 \wedge \xi_4) \vee (\xi_4 \wedge \xi_5). \tag{6.36}$$

It is clear that the truth function is

$$f(x_1, x_2, x_3, x_4, x_5) = \begin{cases} 1, & \text{if } x_1 + x_2 = 2 \\ 1, & \text{if } x_2 + x_3 = 2 \\ 1, & \text{if } x_3 + x_4 = 2 \\ 1, & \text{if } x_4 + x_5 = 2 \\ 0, & \text{otherwise.} \end{cases}$$

A run of the truth value solver shows that $T(X) = 0.7$.

Example 6.14: Assume that $\xi_1, \xi_2, \xi_3, \xi_4, \xi_5$ are independent uncertain propositions with truth values $0.1, 0.3, 0.5, 0.7, 0.9$, respectively. Let

$$X = \text{"only 4 propositions of } \xi_1, \xi_2, \xi_3, \xi_4, \xi_5 \text{ are true"}. \tag{6.37}$$

It is clear that the truth function is

$$f(x_1, x_2, x_3, x_4, x_5) = \begin{cases} 1, & \text{if } x_1 + x_2 + x_3 + x_4 + x_5 = 4 \\ 0, & \text{if } x_1 + x_2 + x_3 + x_4 + x_5 \neq 4. \end{cases}$$

A run of the truth value solver shows that $T(X) = 0.3$.

Example 6.15: Assume that $\xi_1, \xi_2, \xi_3, \xi_4, \xi_5$ are independent uncertain propositions with truth values $0.1, 0.3, 0.5, 0.7, 0.9$, respectively. Let

$$X = \text{"only odd number of propositions of } \xi_1, \xi_2, \xi_3, \xi_4, \xi_5 \text{ are true"}. \tag{6.38}$$

It is clear that the truth function is

$$f(x_1, x_2, x_3, x_4, x_5) = \begin{cases} 1, & \text{if } x_1 + x_2 + x_3 + x_4 + x_5 \in \{1, 3, 5\} \\ 0, & \text{if } x_1 + x_2 + x_3 + x_4 + x_5 \in \{0, 2, 4\}. \end{cases}$$

A run of the truth value solver shows that $T(X) = 0.5$.

Chapter 7

Uncertain Entailment

Uncertain entailment, developed by Liu [131] in 2009, is a methodology for calculating the truth value of an uncertain formula via the maximum uncertainty principle when the truth values of other uncertain formulas are given. In order to solve this problem, this chapter will introduce an entailment model. As applications of uncertain entailment, this chapter will also discuss modus ponens, modus tollens, and hypothetical syllogism.

7.1 Entailment Model

Assume $\xi_1, \xi_2, \cdots, \xi_n$ are independent uncertain propositions with *unknown* truth values $\alpha_1, \alpha_2, \cdots, \alpha_n$, respectively. Also assume that $X_1, X_2, \cdots, X_m$ are uncertain formulas containing $\xi_1, \xi_2, \cdots, \xi_n$ with *known* truth values β_1, $\beta_2, \cdots, \beta_m$, respectively. Now let X be an additional uncertain formula containing $\xi_1, \xi_2, \cdots, \xi_n$. What is the truth value of X?

This is just the uncertain entailment problem. In order to solve it, let us consider what values $\alpha_1, \alpha_2, \cdots, \alpha_n$ may take. The first constraint is

$$0 \le \alpha_j \le 1, \quad j = 1, 2, \cdots, n. \tag{7.1}$$

We also hope

$$T(X_i) = \beta_i, \quad i = 1, 2, \cdots, m \tag{7.2}$$

where each $T(X_i)$ ($1 \le i \le m$) is determined by the truth function f_i as follows,

$$T(X_i) = \begin{cases} \displaystyle\sup_{f_i(x_1, x_2, \cdots, x_n)=1} \min_{1 \le j \le n} \nu_j(x_j), \\ \qquad \text{if } \displaystyle\sup_{f_i(x_1, x_2, \cdots, x_n)=1} \min_{1 \le j \le n} \nu_j(x_j) < 0.5 \\ 1 - \displaystyle\sup_{f_i(x_1, x_2, \cdots, x_n)=0} \min_{1 \le j \le n} \nu_j(x_j), \\ \qquad \text{if } \displaystyle\sup_{f_i(x_1, x_2, \cdots, x_n)=1} \min_{1 \le j \le n} \nu_j(x_j) \ge 0.5 \end{cases} \tag{7.3}$$

and

$$\nu_j(x_j) = \begin{cases} \alpha_j, & \text{if } x_j = 1 \\ 1 - \alpha_j, & \text{if } x_j = 0 \end{cases} \tag{7.4}$$

for $j = 1, 2, \cdots, n$.

Based on the truth values $\alpha_1, \alpha_2, \cdots, \alpha_n$ and truth function f, the truth value of X is

$$T(X) = \begin{cases} \displaystyle\sup_{f(x_1,x_2,\cdots,x_n)=1} \min_{1 \le j \le n} \nu_j(x_j), \\ \qquad \text{if } \displaystyle\sup_{f(x_1,x_2,\cdots,x_n)=1} \min_{1 \le j \le n} \nu_j(x_j) < 0.5 \\ 1 - \displaystyle\sup_{f(x_1,x_2,\cdots,x_n)=0} \min_{1 \le j \le n} \nu_j(x_j), \\ \qquad \text{if } \displaystyle\sup_{f(x_1,x_2,\cdots,x_n)=1} \min_{1 \le j \le n} \nu_j(x_j) \ge 0.5. \end{cases} \tag{7.5}$$

Since the truth values $\alpha_1, \alpha_2, \cdots, \alpha_n$ are not uniquely determined, the truth value $T(X)$ is not unique too. For this case, we have to use the maximum uncertainty principle to determine the truth value $T(X)$. That is, $T(X)$ should be assigned the value as close to 0.5 as possible. In other words, we should minimize the value $|T(X) - 0.5|$ via choosing appreciate values of $\alpha_1, \alpha_2, \cdots, \alpha_n$.

In order to find the optimal solution $(\alpha_1, \alpha_2, \cdots, \alpha_n)$, Liu [131] presented the following entailment model,

$$\begin{cases} \min |T(X) - 0.5| \\ \text{subject to:} \\ \quad T(X_i) = \beta_i, \quad i = 1, 2, \cdots, m \\ \quad 0 \le \alpha_j \le 1, \quad j = 1, 2, \cdots, n \end{cases} \tag{7.6}$$

where $T(X_1), T(X_2), \cdots, T(X_m), T(X)$ are functions of $\alpha_1, \alpha_2, \cdots, \alpha_n$ via (7.3) and (7.5).

If the entailment model (7.6) has no feasible solution, then the truth values $\beta_1, \beta_2, \cdots, \beta_m$ are inconsistent with each other. For this case, we cannot entail anything on the uncertain formula X.

If the entailment model (7.6) has an optimal solution $(\alpha_1^*, \alpha_2^*, \cdots, \alpha_n^*)$, then the truth value of X is just (7.5) with

$$\nu_j(x_j) = \begin{cases} \alpha_j^*, & \text{if } x_j = 1 \\ 1 - \alpha_j^*, & \text{if } x_j = 0 \end{cases} \tag{7.7}$$

for $j = 1, 2, \cdots, n$.

Example 7.1: Let ξ_1 and ξ_2 be independent uncertain propositions with unknown truth values α_1 and α_2, respectively. It is known that

$$T(\xi_1 \vee \xi_2) = \beta_1, \quad T(\xi_1 \wedge \xi_2) = \beta_2. \tag{7.8}$$

What is the truth value of $\xi_1 \to \xi_2$? In order to answer this question, we write

$$X_1 = \xi_1 \vee \xi_2, \quad X_2 = \xi_1 \wedge \xi_2, \quad X = \xi_1 \to \xi_2.$$

Then we have

$$T(X_1) = \alpha_1 \vee \alpha_2 = \beta_1,$$

$$T(X_2) = \alpha_1 \wedge \alpha_2 = \beta_2,$$

$$T(X) = (1 - \alpha_1) \vee \alpha_2.$$

For this case, the entailment model (7.6) becomes

$$\begin{cases} \min |(1 - \alpha_1) \vee \alpha_2 - 0.5| \\ \text{subject to:} \\ \quad \alpha_1 \vee \alpha_2 = \beta_1 \\ \quad \alpha_1 \wedge \alpha_2 = \beta_2 \\ \quad 0 \leq \alpha_1 \leq 1 \\ \quad 0 \leq \alpha_2 \leq 1. \end{cases} \tag{7.9}$$

When $\beta_1 \geq \beta_2$, there are only two feasible solutions $(\alpha_1, \alpha_2) = (\beta_1, \beta_2)$ and $(\alpha_1, \alpha_2) = (\beta_2, \beta_1)$. If $\beta_1 + \beta_2 < 1$, the optimal solution produces

$$T(X) = (1 - \alpha_1^*) \vee \alpha_2^* = 1 - \beta_1;$$

if $\beta_1 + \beta_2 = 1$, the optimal solution produces

$$T(X) = (1 - \alpha_1^*) \vee \alpha_2^* = \beta_1 \text{ or } \beta_2;$$

if $\beta_1 + \beta_2 > 1$, the optimal solution produces

$$T(X) = (1 - \alpha_1^*) \vee \alpha_2^* = \beta_2.$$

When $\beta_1 < \beta_2$, there is no feasible solution and the truth values are ill-assigned. As a summary, we have

$$T(\xi_1 \to \xi_2) = \begin{cases} 1 - \beta_1, & \text{if } \beta_1 \geq \beta_2 \text{ and } \beta_1 + \beta_2 < 1 \\ \beta_1 \text{ or } \beta_2, & \text{if } \beta_1 \geq \beta_2 \text{ and } \beta_1 + \beta_2 = 1 \\ \beta_2, & \text{if } \beta_1 \geq \beta_2 \text{ and } \beta_1 + \beta_2 > 1 \\ \text{illness}, & \text{if } \beta_1 < \beta_2. \end{cases} \tag{7.10}$$

Example 7.2: Let ξ_1, ξ_2, ξ_3 be independent uncertain propositions with unknown truth values $\alpha_1, \alpha_2, \alpha_3$, respectively. It is known that

$$T(\xi_1 \to \xi_2) = \beta_1, \quad T(\xi_2 \to \xi_3) = \beta_2. \tag{7.11}$$

What is the truth value of ξ_2? In order to answer this question, we write

$$X_1 = \xi_1 \to \xi_2, \quad X_2 = \xi_2 \to \xi_3, \quad X = \xi_2.$$

Then we have

$$T(X_1) = (1 - \alpha_1) \vee \alpha_2 = \beta_1,$$

$$T(X_2) = (1 - \alpha_2) \vee \alpha_3 = \beta_2,$$

$$T(X) = \alpha_2.$$

For this case, the entailment model (7.6) becomes

$$\begin{cases} \min |\alpha_2 - 0.5| \\ \text{subject to:} \\ \quad (1 - \alpha_1) \vee \alpha_2 = \beta_1 \\ \quad (1 - \alpha_2) \vee \alpha_3 = \beta_2 \\ \quad 0 \leq \alpha_1 \leq 1 \\ \quad 0 \leq \alpha_2 \leq 1 \\ \quad 0 \leq \alpha_3 \leq 1. \end{cases} \tag{7.12}$$

The optimal solution $(\alpha_1^*, \alpha_2^*, \alpha_3^*)$ produces

$$T(\xi_2) = \begin{cases} \beta_1, & \text{if } \beta_1 + \beta_2 \geq 1 \text{ and } \beta_1 < 0.5 \\ 1 - \beta_2, & \text{if } \beta_1 + \beta_2 \geq 1 \text{ and } \beta_2 < 0.5 \\ 0.5, & \text{if } \beta_1 \geq 0.5 \text{ and } \beta_2 \geq 0.5 \\ \text{illness}, & \text{if } \beta_1 + \beta_2 < 1. \end{cases} \tag{7.13}$$

Example 7.3: Let ξ_1, ξ_2, ξ_3 be independent uncertain propositions with unknown truth values $\alpha_1, \alpha_2, \alpha_3$, respectively. It is known that

$$T(\xi_1 \to \xi_2) = \beta_1, \quad T(\xi_1 \to \xi_3) = \beta_2. \tag{7.14}$$

What is the truth value of $\xi_1 \to \xi_2 \wedge \xi_3$? In order to answer this question, we write

$$X_1 = \xi_1 \to \xi_2, \quad X_2 = \xi_1 \to \xi_3, \quad X = \xi_1 \to \xi_2 \wedge \xi_3.$$

Then we have

$$T(X_1) = (1 - \alpha_1) \vee \alpha_2 = \beta_1,$$

$$T(X_2) = (1 - \alpha_1) \vee \alpha_3 = \beta_2,$$

$$T(X) = (1 - \alpha_1) \vee (\alpha_2 \wedge \alpha_3).$$

For this case, the entailment model (7.6) becomes

$$
\begin{cases}
\min |(1 - \alpha_1) \vee (\alpha_2 \wedge \alpha_3) - 0.5| \\
\text{subject to:} \\
\quad (1 - \alpha_1) \vee \alpha_2 = \beta_1 \\
\quad (1 - \alpha_1) \vee \alpha_3 = \beta_2 \\
\quad 0 \le \alpha_1 \le 1 \\
\quad 0 \le \alpha_2 \le 1 \\
\quad 0 \le \alpha_3 \le 1.
\end{cases}
\qquad (7.15)
$$

The optimal solution $(\alpha_1^*, \alpha_2^*, \alpha_3^*)$ produces $T(\xi_1 \to \xi_2 \wedge \xi_3) = \beta_1 \wedge \beta_2$.

Example 7.4: Let ξ_1, ξ_2, ξ_3 be independent uncertain propositions with unknown truth values $\alpha_1, \alpha_2, \alpha_3$, respectively. It is known that

$$
T(\xi_1 \to \xi_2) = \beta_1, \quad T(\xi_1 \to \xi_3) = \beta_2. \qquad (7.16)
$$

What is the truth value of $\xi_1 \to \xi_2 \vee \xi_3$? In order to answer this question, we write

$$
X_1 = \xi_1 \to \xi_2, \quad X_2 = \xi_1 \to \xi_3, \quad X = \xi_1 \to \xi_2 \vee \xi_3.
$$

Then we have

$$
T(X_1) = (1 - \alpha_1) \vee \alpha_2 = \beta_1,
$$

$$
T(X_2) = (1 - \alpha_1) \vee \alpha_3 = \beta_2,
$$

$$
T(X) = (1 - \alpha_1) \vee \alpha_2 \vee \alpha_3.
$$

For this case, the entailment model (7.6) becomes

$$
\begin{cases}
\min |(1 - \alpha_1) \vee \alpha_2 \vee \alpha_3 - 0.5| \\
\text{subject to:} \\
\quad (1 - \alpha_1) \vee \alpha_2 = \beta_1 \\
\quad (1 - \alpha_1) \vee \alpha_3 = \beta_2 \\
\quad 0 \le \alpha_1 \le 1 \\
\quad 0 \le \alpha_2 \le 1 \\
\quad 0 \le \alpha_3 \le 1.
\end{cases}
\qquad (7.17)
$$

The optimal solution $(\alpha_1^*, \alpha_2^*, \alpha_3^*)$ produces $T(\xi_1 \to \xi_2 \vee \xi_3) = \beta_1 \vee \beta_2$.

Example 7.5: Let ξ_1, ξ_2, ξ_3 be independent uncertain propositions with unknown truth values $\alpha_1, \alpha_2, \alpha_3$, respectively. It is known that

$$
T(\xi_1 \to \xi_3) = \beta_1, \quad T(\xi_2 \to \xi_3) = \beta_2. \qquad (7.18)
$$

What is the truth value of $\xi_1 \vee \xi_2 \to \xi_3$? In order to answer this question, we write

$$X_1 = \xi_1 \to \xi_3, \quad X_2 = \xi_2 \to \xi_3, \quad X = \xi_1 \vee \xi_2 \to \xi_3.$$

Then we have

$$T(X_1) = (1 - \alpha_1) \vee \alpha_3 = \beta_1,$$
$$T(X_2) = (1 - \alpha_2) \vee \alpha_3 = \beta_2,$$
$$T(X) = (1 - \alpha_1 \vee \alpha_2) \vee \alpha_3.$$

For this case, the entailment model (7.6) becomes

$$\begin{cases} \min |(1 - \alpha_1 \vee \alpha_2) \vee \alpha_3 - 0.5| \\ \text{subject to:} \\ \quad (1 - \alpha_1) \vee \alpha_3 = \beta_1 \\ \quad (1 - \alpha_2) \vee \alpha_3 = \beta_2 \\ \quad 0 \le \alpha_1 \le 1 \\ \quad 0 \le \alpha_2 \le 1 \\ \quad 0 \le \alpha_3 \le 1. \end{cases} \qquad (7.19)$$

The optimal solution $(\alpha_1^*, \alpha_2^*, \alpha_3^*)$ produces $T(\xi_1 \vee \xi_2 \to \xi_3) = \beta_1 \wedge \beta_2$.

Example 7.6: Let ξ_1, ξ_2, ξ_3 be independent uncertain propositions with unknown truth values $\alpha_1, \alpha_2, \alpha_3$, respectively. It is known that

$$T(\xi_1 \to \xi_3) = \beta_1, \quad T(\xi_2 \to \xi_3) = \beta_2. \qquad (7.20)$$

What is the truth value of $\xi_1 \wedge \xi_2 \to \xi_3$? In order to answer this question, we write

$$X_1 = \xi_1 \to \xi_3, \quad X_2 = \xi_2 \to \xi_3, \quad X = \xi_1 \wedge \xi_2 \to \xi_3.$$

Then we have

$$T(X_1) = (1 - \alpha_1) \vee \alpha_3 = \beta_1,$$
$$T(X_2) = (1 - \alpha_2) \vee \alpha_3 = \beta_2,$$
$$T(X) = (1 - \alpha_1 \wedge \alpha_2) \vee \alpha_3.$$

For this case, the entailment model (7.6) becomes

$$\begin{cases} \min |(1 - \alpha_1 \wedge \alpha_2) \vee \alpha_3 - 0.5| \\ \text{subject to:} \\ \quad (1 - \alpha_1) \vee \alpha_3 = \beta_1 \\ \quad (1 - \alpha_2) \vee \alpha_3 = \beta_2 \\ \quad 0 \le \alpha_1 \le 1 \\ \quad 0 \le \alpha_2 \le 1 \\ \quad 0 \le \alpha_3 \le 1. \end{cases} \qquad (7.21)$$

The optimal solution $(\alpha_1^*, \alpha_2^*, \alpha_3^*)$ produces $T(\xi_1 \wedge \xi_2 \to \xi_3) = \beta_1 \vee \beta_2$.

7.2   Modus Ponens

Classical modus ponens tells us that if both ξ and $\xi \rightarrow \eta$ are true, then η is true. This section provides a version of modus ponens in the framework of uncertain logic.

Theorem 7.1 *(Liu [131], Modus Ponens) Let ξ and η be independent uncertain propositions. Suppose ξ and $\xi \rightarrow \eta$ are two uncertain formulas with truth values β_1 and β_2, respectively. Then the truth value of η is*

$$T(\eta) = \begin{cases} \beta_2, & \text{if } \beta_1 + \beta_2 > 1 \\ 0.5 \wedge \beta_2, & \text{if } \beta_1 + \beta_2 = 1 \\ \text{illness}, & \text{if } \beta_1 + \beta_2 < 1. \end{cases} \qquad (7.22)$$
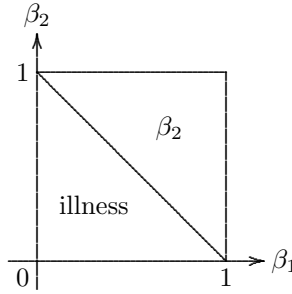
Figure 7.1: Modus Ponens

Proof: Denote the truth values of ξ and η by α_1 and α_2, respectively, and write

$$X_1 = \xi, \quad X_2 = \xi \rightarrow \eta, \quad X = \eta.$$

It is clear that

$$T(X_1) = \alpha_1 = \beta_1,$$

$$T(X_2) = (1 - \alpha_1) \vee \alpha_2 = \beta_2,$$

$$T(X) = \alpha_2.$$

For this case, the entailment model (7.6) becomes

$$\begin{cases} \min |\alpha_2 - 0.5| \\ \text{subject to:} \\ \quad \alpha_1 = \beta_1 \\ \quad (1 - \alpha_1) \vee \alpha_2 = \beta_2 \\ \quad 0 \leq \alpha_1 \leq 1 \\ \quad 0 \leq \alpha_2 \leq 1. \end{cases} \qquad (7.23)$$

When $\beta_1 + \beta_2 > 1$, there is only one feasible solution and then the optimal solution is

$$\alpha_1^* = \beta_1, \quad \alpha_2^* = \beta_2.$$

Thus $T(\eta) = \alpha_2^* = \beta_2$. When $\beta_1 + \beta_2 = 1$, the feasible set is $\{\beta_1\} \times [0, \beta_2]$ and the optimal solution is

$$\alpha_1^* = \beta_1, \quad \alpha_2^* = 0.5 \wedge \beta_2.$$

Thus $T(\eta) = \alpha_2^* = 0.5 \wedge \beta_2$. When $\beta_1 + \beta_2 < 1$, there is no feasible solution and the truth values are ill-assigned. The theorem is proved.

Remark 7.1: Different from the classical logic, the uncertain propositions ξ and η in $\xi \to \eta$ are statements with some truth values rather than pure statements. Thus the truth value of $\xi \to \eta$ is understood as

$$T(\xi \to \eta) = (1 - T(\xi)) \vee T(\eta). \tag{7.24}$$

Remark 7.2: Note that $T(\eta)$ in (7.22) does not necessarily represent the objective truth degree of η. For example, if $T(\xi)$ is small, then $T(\eta)$ is the truth value that η might (not must) be true.

7.3 Modus Tollens

Classical modus tollens tells us that if $\xi \to \eta$ is true and η is false, then ξ is false. This section provides a version of modus tollens in the framework of uncertain logic.

Theorem 7.2 *(Liu [131], Modus Tollens) Let ξ and η be independent uncertain propositions. Suppose $\xi \to \eta$ and η are two uncertain formulas with truth values β_1 and β_2, respectively. Then the truth value of ξ is*

$$T(\xi) = \begin{cases} 1 - \beta_1, & \text{if } \beta_1 > \beta_2 \\ (1 - \beta_1) \vee 0.5, & \text{if } \beta_1 = \beta_2 \\ \text{illness}, & \text{if } \beta_1 < \beta_2. \end{cases} \tag{7.25}$$

Proof: Denote the truth values of ξ and η by α_1 and α_2, respectively, and write

$$X_1 = \xi \to \eta, \quad X_2 = \eta, \quad X = \xi.$$

It is clear that

$$T(X_1) = (1 - \alpha_1) \vee \alpha_2 = \beta_1,$$

$$T(X_2) = \alpha_2 = \beta_2,$$

$$T(X) = \alpha_1.$$
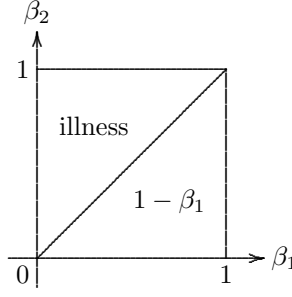
Figure 7.2: Modus Tollens

For this case, the entailment model (7.6) becomes

$$
\begin{cases}
\min |\alpha_1 - 0.5| \\
\quad \text{subject to:} \\
\qquad (1 - \alpha_1) \vee \alpha_2 = \beta_1 \\
\qquad \alpha_2 = \beta_2 \\
\qquad 0 \le \alpha_1 \le 1 \\
\qquad 0 \le \alpha_2 \le 1.
\end{cases}
\tag{7.26}
$$

When $\beta_1 > \beta_2$, there is only one feasible solution and then the optimal solution is

$$
\alpha_1^* = 1 - \beta_1, \quad \alpha_2^* = \beta_2.
$$

Thus $T(\xi) = \alpha_1^* = 1 - \beta_1$. When $\beta_1 = \beta_2$, the feasible set is $[1 - \beta_1, 1] \times \{\beta_2\}$ and the optimal solution is

$$
\alpha_1^* = (1 - \beta_1) \vee 0.5, \quad \alpha_2^* = \beta_2.
$$

Thus $T(\xi) = \alpha_1^* = (1 - \beta_1) \vee 0.5$. When $\beta_1 < \beta_2$, there is no feasible solution and the truth values are ill-assigned. The theorem is proved.

7.4 Hypothetical Syllogism

Classical hypothetical syllogism tells us that if both $\xi \to \eta$ and $\eta \to \tau$ are true, then $\xi \to \tau$ is true. This section provides a version of hypothetical syllogism in the framework of uncertain logic.

Theorem 7.3 *(Liu [131], Hypothetical Syllogism) Let ξ, η, τ be independent uncertain propositions. Suppose $\xi \to \eta$ and $\eta \to \tau$ are two uncertain formulas with truth values β_1 and β_2, respectively. Then the truth value of $\xi \to \tau$ is*

$$
T(\xi \to \tau) = \begin{cases}
\beta_1 \wedge \beta_2, & \text{if } \beta_1 \wedge \beta_2 \ge 0.5 \\
0.5, & \text{if } \beta_1 + \beta_2 \ge 1 \text{ and } \beta_1 \wedge \beta_2 < 0.5 \\
\text{illness}, & \text{if } \beta_1 + \beta_2 < 1.
\end{cases}
\tag{7.27}
$$
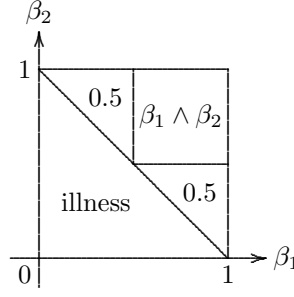
Figure 7.3: Hypothetical Syllogism

Proof: Denote the truth values of ξ, η, τ by $\alpha_1, \alpha_2, \alpha_3$, respectively, and write

$$X_1 = \xi \rightarrow \eta, \quad X_2 = \eta \rightarrow \tau, \quad X = \xi \rightarrow \tau.$$

It is clear that

$$T(X_1) = (1 - \alpha_1) \vee \alpha_2 = \beta_1,$$
$$T(X_2) = (1 - \alpha_2) \vee \alpha_3 = \beta_2,$$
$$T(X) = (1 - \alpha_1) \vee \alpha_3.$$

For this case, the entailment model (7.6) becomes

$$
\begin{cases}
\min |(1 - \alpha_1) \vee \alpha_3 - 0.5| \\
\text{subject to:} \\
\quad (1 - \alpha_1) \vee \alpha_2 = \beta_1 \\
\quad (1 - \alpha_2) \vee \alpha_3 = \beta_2 \\
\quad 0 \leq \alpha_1 \leq 1 \\
\quad 0 \leq \alpha_2 \leq 1 \\
\quad 0 \leq \alpha_3 \leq 1.
\end{cases}
\tag{7.28}
$$

When $\beta_1 \wedge \beta_2 \geq 0.5$, we have

$$T(\xi \rightarrow \tau) = (1 - \alpha_1^*) \vee \alpha_3^* = \beta_1 \wedge \beta_2.$$

When $\beta_1 + \beta_2 \geq 1$ and $\beta_1 \wedge \beta_2 < 0.5$, we have

$$T(\xi \rightarrow \tau) = (1 - \alpha_1^*) \vee \alpha_3^* = 0.5.$$

When $\beta_1 + \beta_2 < 1$, there is no feasible solution and the truth values are ill-assigned. The theorem is proved.

7.5   Automatic Entailment Machine

Automatic Entailment Machine is a software for solving the entailment model. This software may be downloaded from http://orsc.edu.cn/liu/resources.htm.

Chapter 8

Uncertain Inference

Uncertain inference was proposed by Liu [130] in 2009 as a process of deriving consequences from uncertain knowledge or evidence via the tool of conditional uncertainty. This chapter will introduce the concepts of uncertain sets and matching degree, and provide some uncertain inference rules. Finally, uncertain system and uncertain inference control are presented by the tool of uncertain inference.

8.1   Uncertain Set

Before introducing uncertain inference, let us recall the concept of uncertain set proposed by Liu [127] in 2007.

Definition 8.1 *(Liu [127]) An uncertain set is a measurable function ξ from an uncertainty space $(\Gamma, \mathcal{L}, \mathcal{M})$ to a collection of sets of real numbers, i.e., for any Borel set B, the set*

$$\{\gamma \in \Gamma \mid \xi(\gamma) \cap B \neq \emptyset\} \tag{8.1}$$

is an event.

Example 8.1: Let $(\Gamma, \mathcal{L}, \mathcal{M})$ be an uncertainty space with $\Gamma = \{\gamma_1, \gamma_2, \gamma_3\}$. Then the function

$$\xi(\gamma) = \begin{cases} [1,3], & \text{if } \gamma = \gamma_1 \\ [2,4], & \text{if } \gamma = \gamma_2 \\ [3,5], & \text{if } \gamma = \gamma_3 \end{cases}$$

is an uncertain set on $(\Gamma, \mathcal{L}, \mathcal{M})$.

Example 8.2: Let $(\Gamma, \mathcal{L}, \mathcal{M})$ be an uncertainty space. Then the function

$$\xi(\gamma) = [\gamma, \gamma + 1], \quad \forall \gamma \in \Gamma$$

is an uncertain set on $(\Gamma, \mathcal{L}, \mathcal{M})$.

Example 8.3: Let $(\Gamma, \mathcal{L}, \mathcal{M})$ be an uncertainty space with $\Gamma = [0, +\infty)$. Then the function

$$\xi(\gamma) = \left[-\frac{1}{1 + \gamma^2}, \frac{1}{1 + \gamma^2} \right], \quad \forall \gamma \in \Gamma$$

is an uncertain set on $(\Gamma, \mathcal{L}, \mathcal{M})$.

Definition 8.2 *Let ξ and η be two uncertain sets on the uncertainty space $(\Gamma, \mathcal{L}, \mathcal{M})$. Then the complement ξ^c of uncertain set ξ is*

$$\xi^c(\gamma) = (\xi(\gamma))^c, \quad \forall \gamma \in \Gamma. \tag{8.2}$$

The union $\xi \cup \eta$ of uncertain sets ξ and η is

$$(\xi \cup \eta)(\gamma) = \xi(\gamma) \cup \eta(\gamma), \quad \forall \gamma \in \Gamma. \tag{8.3}$$

The intersection $\xi \cap \eta$ of uncertain sets ξ and η is

$$(\xi \cap \eta)(\gamma) = \xi(\gamma) \cap \eta(\gamma), \quad \forall \gamma \in \Gamma. \tag{8.4}$$

It is clear that

$$\xi \cup \xi^c \equiv \Re, \quad \xi \cap \xi^c \equiv \emptyset \tag{8.5}$$

for any uncertain set ξ.

Definition 8.3 *An uncertain set ξ on $(\Gamma, \mathcal{L}, \mathcal{M})$ is said to have a first identification function $\lambda(x)$ if*

$$0 \le \lambda(x) \le \sup_{\gamma \in \Gamma} \mathcal{M}\{\gamma\}, \quad \forall x \in \Re, \tag{8.6}$$

$$\xi(\gamma) = \left\{ x \in \Re \mid \lambda(x) \ge \mathcal{M}\{\gamma\} \right\}, \quad \forall \gamma \in \Gamma. \tag{8.7}$$

It is not true that any uncertain set has its own first identification function. A necessary condition is that $\xi(\gamma_1) \subset \xi(\gamma_2)$ whenever $\mathcal{M}\{\gamma_1\} \ge \mathcal{M}\{\gamma_2\}$. In addition, a first identification function must meet the conditions

$$0 \le \lambda(x) \le 1, \quad \sup_{x \ne y} (\lambda(x) + \lambda(y)) = 1.$$

Definition 8.4 *Let ξ be an uncertain set on $(\Gamma, \mathcal{L}, \mathcal{M})$. A conditional uncertain set of ξ given B is a measurable function $\xi|_B$ from the conditional uncertainty space $(\Gamma, \mathcal{L}, \mathcal{M}\{\cdot|B\})$ to a collection of sets of real numbers such that*

$$\xi|_B(\gamma) \equiv \xi(\gamma), \quad \forall \gamma \in \Gamma \tag{8.8}$$

in the sense of classical set theory.
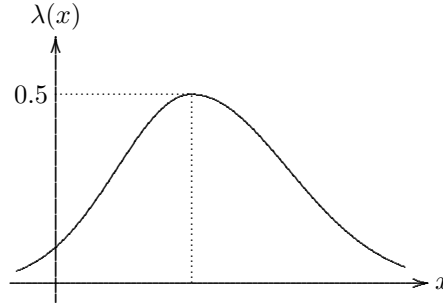
Figure 8.1: A First Identification Function

Definition 8.5 *Let ξ be an uncertain set with a first identification function λ. If λ is a unimodal function about x_0 (i.e., increasing on $(-\infty, x_0)$ and decreasing on $(x_0, +\infty)$), then the expected value of ξ is*

$$E[\xi] = x_0 + \int_{x_0}^{+\infty} \lambda(x)\mathrm{d}x - \int_{-\infty}^{x_0} \lambda(x)\mathrm{d}x \qquad (8.9)$$

provided that at least one of the above integrals is finite.

Example 8.4: By a *rectangular uncertain set* we mean the uncertain set fully determined by the pair (a, b) of crisp numbers with $a < b$, whose first identification function is

$$\lambda(x) = 0.5, \quad a \le x \le b.$$

The rectangular uncertain set $\xi = (a, b)$ has an expected value

$$E[\xi] = \frac{a+b}{2}.$$

Example 8.5: By a *triangular uncertain set* we mean the uncertain set fully determined by the triplet (a, b, c) of crisp numbers with $a < b < c$, whose first identification function is

$$\lambda(x) = \begin{cases} \dfrac{x-a}{2(b-a)}, & \text{if } a \le x \le b \\[2mm] \dfrac{x-c}{2(b-c)}, & \text{if } b \le x \le c. \end{cases}$$

The rectangular uncertain set $\xi = (a, b, c)$ has an expected value

$$E[\xi] = \frac{a + 2b + c}{4}.$$

Example 8.6: By a *trapezoidal uncertain set* we mean the uncertain set fully determined by the quadruplet (a, b, c, d) of crisp numbers with $a < b < c < d$, whose first identification function is

$$\lambda(x) = \begin{cases} \dfrac{x - a}{2(b - a)}, & \text{if } a \leq x \leq b \\[2mm] 0.5, & \text{if } b \leq x \leq c \\[2mm] \dfrac{x - d}{2(c - d)}, & \text{if } c \leq x \leq d. \end{cases}$$

The trapezoidal uncertain set $\xi = (a, b, c, d)$ has an expected value

$$E[\xi] = \frac{a + b + c + d}{4}.$$

8.2   Matching Degree

Let ξ and ξ^* be two uncertain sets. In order to define the matching degree that ξ^* matches ξ, we first introduce the following symbols:

$$\{\xi^* \subset \xi\} = \{\gamma \in \Gamma \mid \xi^*(\gamma) \subset \xi(\gamma)\}, \tag{8.10}$$

$$\{\xi^* \subset \xi^c\} = \{\gamma \in \Gamma \mid \xi^*(\gamma) \subset \xi(\gamma)^c\}, \tag{8.11}$$

$$\{\xi^* \not\subset \xi^c\} = \{\gamma \in \Gamma \mid \xi^*(\gamma) \cap \xi(\gamma) \neq \emptyset\}. \tag{8.12}$$
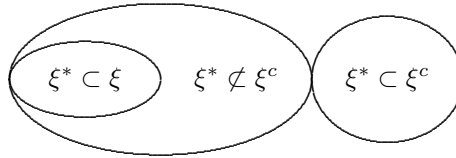
Figure 8.2: Inclusion Relationship between ξ and ξ^*

We would like to know the matching degree that ξ^* is included in ξ. Note that we cannot use the symbol "$\subset$" to represent the inclusion of uncertain sets because the symbol has different meaning in classical set theory. Thus we need a new symbol "$\xi^* \rhd \xi$" to represent the inclusion relation that reads as "ξ^* matches ξ". An unassailable fact is

$$\{\xi^* \subset \xi\} \subset \{\xi^* \rhd \xi\} \subset \{\xi^* \not\subset \xi^c\}. \tag{8.13}$$

How do we determine the event $\xi^* \rhd \xi$? The following definition answers this question.

Definition 8.6 *Let ξ and ξ^* be two uncertain sets. Then the event $\xi^* \rhd \xi$ (i.e., ξ^* matches ξ) is defined as*

$$\{\xi^* \rhd \xi\} = \begin{cases} \{\xi^* \subset \xi\}, & \text{if } \mathcal{M}\{\xi^* \subset \xi\} > \mathcal{M}\{\xi^* \subset \xi^c\} \\ \{\xi^* \not\subset \xi^c\}, & \text{if } \mathcal{M}\{\xi^* \subset \xi\} \leq \mathcal{M}\{\xi^* \subset \xi^c\}. \end{cases} \tag{8.14}$$

When $\mathcal{M}\{\xi^* \subset \xi\} > \mathcal{M}\{\xi^* \subset \xi^c\}$, it follows from the self-duality of uncertain measure that $\mathcal{M}\{\xi^* \subset \xi\} > 1 - \mathcal{M}\{\xi^* \not\subset \xi^c\}$, i.e.,

$$0.5 - \mathcal{M}\{\xi^* \subset \xi\} < \mathcal{M}\{\xi^* \not\subset \xi^c\} - 0.5.$$

Thus we take the event $\{\xi^* \subset \xi\}$ because $\mathcal{M}\{\xi^* \subset \xi\}$ is more closed to 0.5 than $\mathcal{M}\{\xi^* \not\subset \xi^c\}$. When $\mathcal{M}\{\xi^* \subset \xi\} \leq \mathcal{M}\{\xi^* \subset \xi^c\}$, it follows from the self-duality of uncertain measure that $\mathcal{M}\{\xi^* \subset \xi\} \leq 1 - \mathcal{M}\{\xi^* \not\subset \xi^c\}$, i.e.,

$$0.5 - \mathcal{M}\{\xi^* \subset \xi\} \geq \mathcal{M}\{\xi^* \not\subset \xi^c\} - 0.5.$$

Thus we take the event $\{\xi^* \not\subset \xi^c\}$ because $\mathcal{M}\{\xi^* \not\subset \xi^c\}$ is more closed to 0.5 than $\mathcal{M}\{\xi^* \subset \xi\}$. That is, we take the event of $\{\xi^* \subset \xi\}$ or $\{\xi^* \not\subset \xi^c\}$ such that its uncertain measure is as closed to 0.5 as possible. This idea coincides with the maximum uncertainty principle.

Definition 8.7 *Let ξ and ξ^* be two uncertain sets. Then the matching degree of $\xi^* \rhd \xi$ is defined as*

$$\mathcal{M}\{\xi^* \rhd \xi\} = \begin{cases} \mathcal{M}\{\xi^* \subset \xi\}, & \text{if } \mathcal{M}\{\xi^* \subset \xi\} > \mathcal{M}\{\xi^* \subset \xi^c\} \\ \mathcal{M}\{\xi^* \not\subset \xi^c\}, & \text{if } \mathcal{M}\{\xi^* \subset \xi\} \leq \mathcal{M}\{\xi^* \subset \xi^c\}. \end{cases} \tag{8.15}$$

Theorem 8.1 *Let ξ and ξ^* be independent uncertain sets with first identification functions λ and λ^*, respectively. If λ and λ^* are continuous functions, then*

$$\mathcal{M}\{\xi^* \rhd \xi\} = \begin{cases} \sup_{x \in \Re} \lambda(x) \wedge \lambda^*(x), & \text{if } \sup_{x \in \Re} \lambda(x) \wedge \lambda^*(x) < 0.5 \\ 1 - \sup_{x \in \Re}\{\lambda(x)|\lambda(x) < \lambda^*(x)\}, & \text{if } \sup_{x \in \Re} \lambda(x) \wedge \lambda^*(x) = 0.5. \end{cases}$$

Proof: The argument breaks down into two cases. Case 1: If

$$\sup_{x \in \Re} \lambda(x) \wedge \lambda^*(x) < 0.5,$$

then we have

$$\mathcal{M}\{\xi^* \not\subset \xi^c\} = \sup_{x \in \Re} \lambda(x) \wedge \lambda^*(x) < 0.5$$

immediately. Since $\{\xi^* \subset \xi\} \subset \{\xi^* \not\subset \xi^c\}$ and $\{\xi^* \not\subset \xi^c\} = \{\xi^* \subset \xi^c\}^c$, we get

$$\mathcal{M}\{\xi^* \subset \xi\} \leq \mathcal{M}\{\xi^* \not\subset \xi^c\} < 0.5 < 1 - \mathcal{M}\{\xi^* \not\subset \xi^c\} = \mathcal{M}\{\xi^* \subset \xi^c\}.$$

Hence by using Definition 8.7, the matching degree is

$$\mathcal{M}\{\xi^* \rhd \xi\} = \mathcal{M}\{\xi^* \not\subset \xi^c\} = \sup_{x \in \Re} \lambda(x) \wedge \lambda^*(x).$$

Case 2: If

$$\sup_{x \in \Re} \lambda(x) \wedge \lambda^*(x) = 0.5,$$

then $\mathcal{M}\{\xi^* \subset \xi^c\} = \mathcal{M}\{\emptyset\} = 0$. Thus $\mathcal{M}\{\xi^* \subset \xi\} \geq \mathcal{M}\{\xi^* \subset \xi^c\}$, and by using Definition 8.7, the matching degree is

$$\mathcal{M}\{\xi^* \rhd \xi\} = \mathcal{M}\{\xi^* \subset \xi\} = 1 - \mathcal{M}\{\xi^* \not\subset \xi\} = 1 - \sup_{x \in \Re} \{\lambda(x) | \lambda(x) < \lambda^*(x)\}.$$

The theorem is proved.

Example 8.7: Let ξ be an uncertain set with first identification function λ, and let a be a constant. If λ is a continuous function, then by using Theorem 8.1, we have

$$\mathcal{M}\{a \rhd \xi\} = \begin{cases} \lambda(a), & \text{if } \lambda(a) < 0.5 \\ 1, & \text{if } \lambda(a) = 0.5. \end{cases}$$

8.3   Inference Rule

Let $\mathbb{X}$ and $\mathbb{Y}$ be two concepts. It is assumed that we only have a rule "if $\mathbb{X}$ is ξ then $\mathbb{Y}$ is η" where ξ and η are two uncertain sets. We first have the following inference rule.

Inference Rule 8.1 *(Liu [130]) Let $\mathbb{X}$ and $\mathbb{Y}$ be two concepts. Assume a rule "if $\mathbb{X}$ is an uncertain set ξ then $\mathbb{Y}$ is an uncertain set η". From $\mathbb{X}$ is an uncertain set ξ^* we infer that $\mathbb{Y}$ is an uncertain set*

$$\eta^* = \eta|_{\xi^* \rhd \xi} \tag{8.16}$$

which is the conditional uncertain set of η given $\xi^* \rhd \xi$. The inference rule is represented by

$$\begin{array}{l} \text{Rule: If } \mathbb{X} \text{ is } \xi \text{ then } \mathbb{Y} \text{ is } \eta \\ \underline{\text{From: } \mathbb{X} \text{ is } \xi^*} \\ \text{Infer: } \mathbb{Y} \text{ is } \eta^* = \eta|_{\xi^* \rhd \xi} \end{array} \tag{8.17}$$

Theorem 8.2 *Let ξ, ξ^* and η be independent uncertain sets with first identification functions λ, λ^* and ν, respectively. If λ, λ^* and ν are continuous functions, then inference rule 8.1 yields that η^* has a first identification function*

$$\nu^*(y) = \begin{cases} \dfrac{\nu(y)}{\sup\limits_{x \in \Re} \lambda(x) \wedge \lambda^*(x)} \wedge 0.5, & \text{if } \sup\limits_{x \in \Re} \lambda(x) \wedge \lambda^*(x) < 0.5 \\[2em] \dfrac{\nu(y)}{1 - \sup\limits_{x \in \Re}\{\lambda(x) | \lambda(x) < \lambda^*(x)\}} \wedge 0.5, & \text{if } \sup\limits_{x \in \Re} \lambda(x) \wedge \lambda^*(x) = 0.5. \end{cases}$$

Proof: On the one hand, from the continuity of λ, λ^* and ν, we immediately have

$$\sup_{x \in \Re} \lambda(x) = \sup_{x \in \Re} \lambda^*(x) = \sup_{y \in \Re} \nu(y) = \sup_{y \in \Re} \nu^*(y) = 0.5.$$

On the other hand, both definition of conditional uncertainty and independence of uncertain sets produce

$$\mathcal{M}\{y \in \eta | \xi^* \rhd \xi\} = \begin{cases} \dfrac{\mathcal{M}\{y \in \eta\}}{\mathcal{M}\{\xi^* \rhd \xi\}}, & \text{if } \dfrac{\mathcal{M}\{y \in \eta\}}{\mathcal{M}\{\xi^* \rhd \xi\}} < 0.5 \\[3mm] 1 - \dfrac{\mathcal{M}\{y \notin \eta\}}{\mathcal{M}\{\xi^* \rhd \xi\}}, & \text{if } \dfrac{\mathcal{M}\{y \notin \eta\}}{\mathcal{M}\{\xi^* \rhd \xi\}} < 0.5 \\[3mm] 0.5, & \text{otherwise.} \end{cases}$$

It follows from inference rule 8.1 that

$$\nu^*(y) = \frac{\mathcal{M}\{y \in \eta\}}{\mathcal{M}\{\xi^* \rhd \xi\}} \wedge 0.5$$

for any $y \in \Re$. By using Theorem 8.1, we have

$$\mathcal{M}\{y \in \eta\} = \begin{cases} \nu(y), & \text{if } \nu(y) < 0.5 \\ 1, & \text{if } \nu(y) = 0.5. \end{cases}$$

Thus for any $y \in \Re$, we have

$$\nu^*(y) = \frac{\nu(y)}{\mathcal{M}\{\xi^* \rhd \xi\}} \wedge 0.5.$$

Now the argument breaks down into two cases. Case 1: If

$$\sup_{x \in \Re} \lambda(x) \wedge \lambda^*(x) < 0.5,$$

then by using Theorem 8.1, the matching degree is

$$\mathcal{M}\{\xi^* \rhd \xi\} = \sup_{x \in \Re} \lambda(x) \wedge \lambda^*(x),$$

and for this case we obtain

$$\nu^*(y) = \frac{\nu(y)}{\sup\limits_{x \in \Re} \lambda(x) \wedge \lambda^*(x)} \wedge 0.5.$$

Case 2: If

$$\sup_{x \in \Re} \lambda(x) \wedge \lambda^*(x) = 0.5,$$

then by using Theorem 8.1, the matching degree is

$$\mathcal{M}\{\xi^* \rhd \xi\} = 1 - \sup_{x \in \Re} \left\{ \lambda(x) | \lambda(x) < \lambda^*(x) \right\}.$$

For this case, we obtain

$$\nu^*(y) = \frac{\nu(y)}{1 - \sup\limits_{x \in \Re} \{\lambda(x) | \lambda(x) < \lambda^*(x)\}} \wedge 0.5.$$

The theorem is proved. Note that if $\sup \lambda(x) \wedge \lambda^*(x) = 0$, then we set $\nu^*(y) \equiv 0.5$. This means that "if nothing is known then anything is possible".
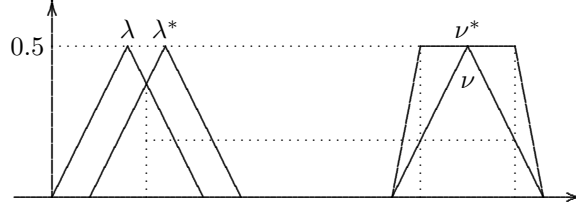
Figure 8.3: Graphical Illustration of Inference Rule

Example 8.8: Let ξ, ξ^* and η be independent uncertain sets with first identification functions λ, λ^* and ν, respectively. If $\lambda^*(x) \leq \lambda(x)$ for all x, then ξ^* matches ξ completely, and inference rule 8.1 yields that η^* has a first identification function

$$\nu^*(y) \equiv \nu(y).$$

Example 8.9: Let ξ and η be independent uncertain sets with first identification functions λ and ν, respectively. If λ and ν are continuous functions and ξ^* is a constant a, then inference rule 8.1 yields that η^* has a first identification function

$$\nu^*(y) = \begin{cases} \dfrac{\nu(y)}{\lambda(a)} \wedge 0.5, & \text{if } \lambda(a) < 0.5 \\[2mm] \nu(y), & \text{if } \lambda(a) = 0.5. \end{cases}$$

8.4   Inference Rule with Multiple Antecedents

Inference Rule 8.2 *(Gao and Ralescu [44]) Let $\mathbb{X}$, $\mathbb{Y}$ and $\mathbb{Z}$ be three concepts. Assume a rule "if $\mathbb{X}$ is an uncertain set ξ and $\mathbb{Y}$ is an uncertain set η then $\mathbb{Z}$ is an uncertain set τ". From $\mathbb{X}$ is an uncertain set ξ^* and $\mathbb{Y}$ is an uncertain set η^* we infer that $\mathbb{Z}$ is an uncertain set*

$$\tau^* = \tau\big|_{(\xi^* \triangleright \xi) \cap (\eta^* \triangleright \eta)} \tag{8.18}$$

which is the conditional uncertain set of τ given $\xi^* \rhd \xi$ and $\eta^* \rhd \eta$. The inference rule is represented by

$$
\begin{array}{l}
\text{Rule: If } \mathbb{X} \text{ is } \xi \text{ and } \mathbb{Y} \text{ is } \eta \text{ then } \mathbb{Z} \text{ is } \tau \\
\underline{\text{From: } \mathbb{X} \text{ is } \xi^* \text{ and } \mathbb{Y} \text{ is } \eta^*} \\
\text{Infer: } \mathbb{Z} \text{ is } \tau^* = \tau|_{(\xi^* \rhd \xi) \cap (\eta^* \rhd \eta)}
\end{array}
\qquad (8.19)
$$

Theorem 8.3 *Let $\xi, \xi^*, \eta, \eta^*, \tau$ be independent uncertain sets with continuous first identification functions $\lambda, \lambda^*, \mu, \mu^*, \nu$, respectively. Then inference rule 8.2 yields that τ^* has a first identification function*

$$
\nu^*(z) = \begin{cases}
\dfrac{\nu(z)}{\sup\limits_{x \in \Re} \lambda(x) \wedge \lambda^*(x) \wedge \sup\limits_{y \in \Re} \mu(y) \wedge \mu^*(y)} \wedge 0.5, \\
\qquad \text{if } \sup\limits_{x \in \Re} \lambda(x) \wedge \lambda^*(x) \wedge \sup\limits_{y \in \Re} \mu(y) \wedge \mu^*(y) < 0.5 \\[2ex]
\dfrac{\nu(z)}{1 - \sup\limits_{x \in \Re} \{\lambda(x)|\lambda(x) < \lambda^*(x)\} \vee \sup\limits_{y \in \Re} \{\mu(y)|\mu(y) < \mu^*(y)\}} \wedge 0.5, \\
\qquad \text{if } \sup\limits_{x \in \Re} \lambda(x) \wedge \lambda^*(x) \wedge \sup\limits_{y \in \Re} \mu(y) \wedge \mu^*(y) = 0.5.
\end{cases}
$$

Proof: It follows from inference rule 8.2 that the uncertain set τ^* has a first identification function

$$
\nu^*(z) = \frac{\mathcal{M}\{z \in \tau\}}{\mathcal{M}\{(\xi^* \rhd \xi) \cap (\eta^* \rhd \eta)\}} \wedge 0.5
$$

for $z \in \Re$. Note that

$$
\mathcal{M}\{z \in \tau\} = \begin{cases}
\nu(z), & \text{if } \nu(z) < 0.5 \\
1, & \text{if } \nu(z) = 0.5.
\end{cases}
$$

Thus for any $z \in \Re$, we have

$$
\nu^*(z) = \frac{\nu(z)}{\mathcal{M}\{(\xi^* \rhd \xi) \cap (\eta^* \rhd \eta)\}} \wedge 0.5.
$$

Now the argument breaks down into two cases. Case 1: If

$$
\sup\limits_{x \in \Re} \lambda(x) \wedge \lambda^*(x) \wedge \sup\limits_{y \in \Re} \mu(y) \wedge \mu^*(y) < 0.5,
$$

then the matching degree is

$$
\mathcal{M}\{(\xi^* \rhd \xi) \cap (\eta^* \rhd \eta)\} = \sup\limits_{x \in \Re} \lambda(x) \wedge \lambda^*(x) \wedge \sup\limits_{y \in \Re} \mu(y) \wedge \mu^*(y),
$$

and for this case we obtain

$$
\nu^*(z) = \frac{\nu(z)}{\sup\limits_{x \in \Re} \lambda(x) \wedge \lambda^*(x) \wedge \sup\limits_{y \in \Re} \mu(y) \wedge \mu^*(y)} \wedge 0.5.
$$

Case 2: If

$$\sup_{x \in \Re} \lambda(x) \wedge \lambda^*(x) \wedge \sup_{y \in \Re} \mu(y) \wedge \mu^*(y) = 0.5,$$

then the matching degree is

$$\mathcal{M}\{(\xi^* \triangleright \xi) \cap (\eta^* \triangleright \eta)\} = 1 - \sup_{x \in \Re} \{\lambda(x) | \lambda(x) < \lambda^*(x)\} \vee \sup_{y \in \Re} \{\mu(y) | \mu(y) < \mu^*(y)\}.$$

For this case, we obtain

$$\nu^*(z) = \frac{\nu(z)}{1 - \sup_{x \in \Re} \{\lambda(x) | \lambda(x) < \lambda^*(x)\} \vee \sup_{y \in \Re} \{\mu(y) | \mu(y) < \mu^*(y)\}} \wedge 0.5.$$

The theorem is proved.

Example 8.10: Let ξ, ξ^*, η, η^* and τ be independent uncertain sets with first identification functions λ, λ^*, μ, μ^* and ν, respectively. If $\lambda^*(x) \leq \lambda(x)$ for all x and $\mu^*(y) \leq \mu(y)$ for all y, then ξ^* matches ξ and η^* matches η completely, and inference rule 8.2 yields that τ^* has a first identification function

$$\nu^*(z) \equiv \nu(z).$$

Example 8.11: Let ξ, η, τ be independent uncertain sets with continuous first identification functions λ, μ, ν, respectively. If ξ^* is a constant a and η^* is a constant b, then inference rule 8.2 yields that τ^* has a first identification function

$$\nu^*(z) = \begin{cases} \dfrac{\nu(z)}{\lambda(a) \wedge \mu(b)} \wedge 0.5, & \text{if } \lambda(a) \wedge \mu(b) < 0.5 \\[2mm] \nu(z), & \text{if } \lambda(a) \wedge \mu(b) = 0.5. \end{cases}$$

8.5   Inference Rule with Multiple If-Then Rules

Inference Rule 8.3 *(Gao [46]) Let $\mathbb{X}$ and $\mathbb{Y}$ be two concepts. Assume two rules "if $\mathbb{X}$ is an uncertain set ξ_1 then $\mathbb{Y}$ is an uncertain set η_1" and "if $\mathbb{X}$ is an uncertain set ξ_2 then $\mathbb{Y}$ is an uncertain set η_2". From $\mathbb{X}$ is an uncertain set ξ^* we infer that $\mathbb{Y}$ is an uncertain set*

$$\eta^* = \frac{\mathcal{M}\{\xi^* \triangleright \xi_1\} \cdot \eta_1|_{\xi^* \triangleright \xi_1}}{\mathcal{M}\{\xi^* \triangleright \xi_1\} + \mathcal{M}\{\xi^* \triangleright \xi_2\}} + \frac{\mathcal{M}\{\xi^* \triangleright \xi_2\} \cdot \eta_2|_{\xi^* \triangleright \xi_2}}{\mathcal{M}\{\xi^* \triangleright \xi_1\} + \mathcal{M}\{\xi^* \triangleright \xi_2\}}. \qquad (8.20)$$

The inference rule is represented by

$$\begin{array}{l} \text{Rule 1: If } \mathbb{X} \text{ is } \xi_1 \text{ then } \mathbb{Y} \text{ is } \eta_1 \\ \text{Rule 2: If } \mathbb{X} \text{ is } \xi_2 \text{ then } \mathbb{Y} \text{ is } \eta_2 \\ \underline{\text{From: } \mathbb{X} \text{ is } \xi^*} \\ \text{Infer: } \mathbb{Y} \text{ is } \eta^* \text{ determined by (8.20)} \end{array} \qquad (8.21)$$

Theorem 8.4 *Let $\xi_1, \xi_2, \eta_1, \eta_2, \xi^*$ be independent uncertain sets with continuous first identification functions $\lambda_1, \lambda_2, \nu_1, \nu_2, \lambda^*$, respectively. Then inference rule 8.3 yields*

$$\eta^* = \frac{c_1}{c_1 + c_2}\eta_1^* + \frac{c_2}{c_1 + c_2}\eta_2^* \tag{8.22}$$

where c_1 and c_2 are constants determined by

$$c_1 = \begin{cases} \sup_{x \in \Re} \lambda_1(x) \wedge \lambda^*(x), & \text{if } \sup_{x \in \Re} \lambda_1(x) \wedge \lambda^*(x) < 0.5 \\ 1 - \sup_{x \in \Re}\{\lambda_1(x)|\lambda_1(x) < \lambda^*(x)\}, & \text{if } \sup_{x \in \Re} \lambda_1(x) \wedge \lambda^*(x) = 0.5, \end{cases}$$

$$c_2 = \begin{cases} \sup_{x \in \Re} \lambda_2(x) \wedge \lambda^*(x), & \text{if } \sup_{x \in \Re} \lambda_2(x) \wedge \lambda^*(x) < 0.5 \\ 1 - \sup_{x \in \Re}\{\lambda_2(x)|\lambda_2(x) < \lambda^*(x)\}, & \text{if } \sup_{x \in \Re} \lambda_2(x) \wedge \lambda^*(x) = 0.5 \end{cases}$$

and η_1^* and η_2^* are uncertain sets whose first identification functions are respectively given by

$$\nu_1^*(z) = \begin{cases} \dfrac{\nu_1(z)}{\sup\limits_{x \in \Re} \lambda_1(x) \wedge \lambda^*(x)} \wedge 0.5, & \text{if } \sup_{x \in \Re} \lambda_1(x) \wedge \lambda^*(x) < 0.5 \\ \dfrac{\nu_1(z)}{1 - \sup\limits_{x \in \Re}\{\lambda_1(x)|\lambda_1(x) < \lambda^*(x)\}} \wedge 0.5, & \text{if } \sup_{x \in \Re} \lambda_1(x) \wedge \lambda^*(x) = 0.5, \end{cases}$$

$$\nu_2^*(z) = \begin{cases} \dfrac{\nu_2(z)}{\sup\limits_{x \in \Re} \lambda_2(x) \wedge \lambda^*(x)} \wedge 0.5, & \text{if } \sup_{x \in \Re} \lambda_2(x) \wedge \lambda^*(x) < 0.5 \\ \dfrac{\nu_2(z)}{1 - \sup\limits_{x \in \Re}\{\lambda_2(x)|\lambda_2(x) < \lambda^*(x)\}} \wedge 0.5, & \text{if } \sup_{x \in \Re} \lambda_2(x) \wedge \lambda^*(x) = 0.5. \end{cases}$$

Proof: It follows from inference rule 8.3 that the uncertain set η^* is just

$$\eta^* = \frac{\mathcal{M}\{\xi^* \rhd \xi_1\} \cdot \eta_1|_{\xi^* \rhd \xi_1}}{\mathcal{M}\{\xi^* \rhd \xi_1\} + \mathcal{M}\{\xi^* \rhd \xi_2\}} + \frac{\mathcal{M}\{\xi^* \rhd \xi_2\} \cdot \eta_2|_{\xi^* \rhd \xi_2}}{\mathcal{M}\{\xi^* \rhd \xi_1\} + \mathcal{M}\{\xi^* \rhd \xi_2\}}.$$

By using Theorem 8.1, the values $\mathcal{M}\{\xi^* \rhd \xi_1\}$ and $\mathcal{M}\{\xi^* \rhd \xi_2\}$ are just c_1 and c_2, respectively. By using Theorem 8.2, the uncertain sets $\eta_1|_{\xi^* \rhd \xi_1}$ and $\eta_2|_{\xi^* \rhd \xi_2}$ have first identification functions ν_1^* and ν_2^*, respectively. Thus the theorem is verified.

Example 8.12: Let $\xi_1, \xi_2, \eta_1, \eta_2$ be independent uncertain sets with continuous first identification functions $\lambda_1, \lambda_2, \nu_1, \nu_2$, respectively, and let ξ^* degenerate to a constant a. Then inference rule 8.3 yields

$$\eta^* = \frac{c_1}{c_1 + c_2}\eta_1^* + \frac{c_2}{c_1 + c_2}\eta_2^* \tag{8.23}$$

where c_1 and c_2 are constants determined by

$$c_1 = \begin{cases} \lambda_1(a), & \text{if } \lambda_1(a) < 0.5 \\ 1, & \text{if } \lambda_1(a) = 0.5, \end{cases} \tag{8.24}$$

$$c_2 = \begin{cases} \lambda_2(a), & \text{if } \lambda_2(a) < 0.5 \\ 1, & \text{if } \lambda_2(a) = 0.5 \end{cases} \tag{8.25}$$

and η_1^* and η_2^* are uncertain sets whose first identification functions are respectively given by

$$\nu_1^*(z) = \begin{cases} \dfrac{\nu_1(z)}{\lambda_1(a)} \wedge 0.5, & \text{if } \lambda_1(a) < 0.5 \\ \nu_1(z), & \text{if } \lambda_1(a) = 0.5, \end{cases} \tag{8.26}$$

$$\nu_2^*(z) = \begin{cases} \dfrac{\nu_2(z)}{\lambda_2(a)} \wedge 0.5, & \text{if } \lambda_2(a) < 0.5 \\ \nu_2(z), & \text{if } \lambda_2(a) = 0.5. \end{cases} \tag{8.27}$$

8.6 General Inference Rule

Inference Rule 8.4 *Let $\mathbb{X}_1, \mathbb{X}_2, \cdots, \mathbb{X}_m$ be concepts. Assume rules "if $\mathbb{X}_1$ is ξ_{i1} and $\cdots$ and $\mathbb{X}_m$ is ξ_{im} then $\mathbb{Y}$ is η_i" for $i = 1, 2, \cdots, k$. From $\mathbb{X}_1$ is ξ_1^* and $\cdots$ and $\mathbb{X}_m$ is ξ_m^* we infer that $\mathbb{Y}$ is an uncertain set*

$$\eta^* = \sum_{i=1}^{k} \frac{c_i \cdot \eta_i \big|_{(\xi_1^* \rhd \xi_{i1}) \cap (\xi_2^* \rhd \xi_{i2}) \cap \cdots \cap (\xi_m^* \rhd \xi_{im})}}{c_1 + c_2 + \cdots + c_k} \tag{8.28}$$

where the coefficients are determined by

$$c_i = \mathcal{M}\left\{ (\xi_1^* \rhd \xi_{i1}) \cap (\xi_2^* \rhd \xi_{i2}) \cap \cdots \cap (\xi_m^* \rhd \xi_{im}) \right\} \tag{8.29}$$

for $i = 1, 2, \cdots, k$. The inference rule is represented by

$$\begin{array}{l} \text{Rule 1: If } \mathbb{X}_1 \text{ is } \xi_{11} \text{ and } \cdots \text{ and } \mathbb{X}_m \text{ is } \xi_{1m} \text{ then } \mathbb{Y} \text{ is } \eta_1 \\ \text{Rule 2: If } \mathbb{X}_1 \text{ is } \xi_{21} \text{ and } \cdots \text{ and } \mathbb{X}_m \text{ is } \xi_{2m} \text{ then } \mathbb{Y} \text{ is } \eta_2 \\ \cdots \\ \text{Rule } k: \text{If } \mathbb{X}_1 \text{ is } \xi_{k1} \text{ and } \cdots \text{ and } \mathbb{X}_m \text{ is } \xi_{km} \text{ then } \mathbb{Y} \text{ is } \eta_k \\ \underline{\text{From: } \mathbb{X}_1 \text{ is } \xi_1^* \text{ and } \cdots \text{ and } \mathbb{X}_m \text{ is } \xi_m^*} \\ \text{Infer: } \mathbb{Y} \text{ is } \eta^* \text{ determined by (8.28)} \end{array} \tag{8.30}$$

Theorem 8.5 *Assume $\xi_{i1}, \xi_{i2}, \cdots, \xi_{im}, \eta_i, \xi_1^*, \xi_2^*, \cdots, \xi_m^*$ are independent uncertain sets with continuous first identification functions $\lambda_{i1}, \lambda_{i2}, \cdots, \lambda_{im}, \nu_i, \lambda_1^*, \lambda_2^*, \cdots, \lambda_m^*$, $i = 1, 2, \cdots, k$, respectively. Then inference rule 8.4 yields*

$$\eta^* = \sum_{i=1}^{k} \frac{c_i \cdot \eta_i^*}{c_1 + c_2 + \cdots + c_k} \tag{8.31}$$

where c_i are constants determined by

$$
c_i = \begin{cases}
\min_{1 \leq l \leq m} \sup_{x \in \Re} \lambda_{il}(x) \wedge \lambda_l^*(x), \\
\quad \text{if } \min_{1 \leq l \leq m} \sup_{x \in \Re} \lambda_{il}(x) \wedge \lambda_l^*(x) < 0.5 \\
1 - \max_{1 \leq l \leq m} \sup_{x \in \Re} \{\lambda_{il}(x) | \lambda_{il}(x) < \lambda_l^*(x)\}, \\
\quad \text{if } \min_{1 \leq l \leq m} \sup_{x \in \Re} \lambda_{il}(x) \wedge \lambda_l^*(x) = 0.5
\end{cases}
\tag{8.32}
$$

and η_i^* are uncertain sets whose first identification functions are given by

$$
\nu_i^*(z) = \begin{cases}
\dfrac{\nu_i(z)}{\min_{1 \leq l \leq m} \sup_{x \in \Re} \lambda_{il}(x) \wedge \lambda_l^*(x)} \wedge 0.5, \\
\quad \text{if } \min_{1 \leq l \leq m} \sup_{x \in \Re} \lambda_{il}(x) \wedge \lambda_l^*(x) < 0.5 \\
\dfrac{\nu_i(z)}{1 - \max_{1 \leq l \leq m} \sup_{x \in \Re} \{\lambda_{il}(x) | \lambda_{il}(x) < \lambda_l^*(x)\}} \wedge 0.5, \\
\quad \text{if } \min_{1 \leq l \leq m} \sup_{x \in \Re} \lambda_{il}(x) \wedge \lambda_l^*(x) = 0.5
\end{cases}
\tag{8.33}
$$

for $i = 1, 2, \cdots, k$, respectively. Note that $\eta_1^*, \eta_2^*, \cdots, \eta_k^*$ are not necessarily independent.

Proof: For each i, since $\xi_1^* \triangleright \xi_{i1}, \xi_2^* \triangleright \xi_{i2}, \cdots, \xi_m^* \triangleright \xi_{im}$ are independent events, we immediately have

$$
\mathcal{M} \left\{ \bigcap_{j=1}^m (\xi_j^* \triangleright \xi_{ij}) \right\} = \min_{1 \leq j \leq m} \mathcal{M}\{\xi_j^* \triangleright \xi_{ij}\}.
$$

From the above equation, we may prove that (8.32) and (8.33) hold. Hence the theorem is verified by inference rule 8.4 immediately.

Example 8.13: Let $\xi_{i1}, \xi_{i2}, \cdots, \xi_{im}, \eta_i$ be independent uncertain sets with continuous first identification functions $\lambda_{i1}, \lambda_{i2}, \cdots, \lambda_{im}, \nu_i, i = 1, 2, \cdots, k$, respectively. If $\xi_1^*, \xi_2^*, \cdots, \xi_m^*$ degenerate to constants $a_1, a_2, \cdots, a_m$, respectively, then inference rule 8.4 yields

$$
\eta^* = \sum_{i=1}^k \frac{c_i \cdot \eta_i^*}{c_1 + c_2 + \cdots + c_k}
\tag{8.34}
$$

where c_i are constants determined by

$$
c_i = \begin{cases}
\min_{1 \leq l \leq m} \lambda_{il}(a_l), & \text{if } \min_{1 \leq l \leq m} \lambda_{il}(a_l) < 0.5 \\
1, & \text{if } \min_{1 \leq l \leq m} \lambda_{il}(a_l) = 0.5
\end{cases}
\tag{8.35}
$$

and η_i^* are uncertain sets whose first identification functions are given by

$$
\nu_i^*(z) = \begin{cases} \dfrac{\nu_i(z)}{\min\limits_{1 \le l \le m} \lambda_{il}(a_l)} \wedge 0.5, & \text{if } \min\limits_{1 \le l \le m} \lambda_{il}(a_l) < 0.5 \\ \nu_i(z), & \text{if } \min\limits_{1 \le l \le m} \lambda_{il}(a_l) = 0.5 \end{cases} \tag{8.36}
$$

for $i = 1, 2, \cdots, k$, respectively. For this case, the uncertain sets $\eta_1^*, \eta_2^*, \cdots, \eta_k^*$ are independent.

8.7 Uncertain System

By an *uncertain system* we mean a function from its inputs to outputs based on uncertain inference rule. For simplicity, we assume there are m deterministic inputs $\alpha_1, \alpha_2, \cdots, \alpha_m$, and n deterministic outputs $\beta_1, \beta_2, \cdots, \beta_n$. At first, we infer n uncertain sets $\eta_1^*, \eta_2^*, \cdots, \eta_n^*$ from the m deterministic inputs by the rule-base (i.e., a set of if-then rules),

> If ξ_{11} and ξ_{12} and$\cdots$and ξ_{1m} then η_{11} and η_{12} and$\cdots$and η_{1n}
> If ξ_{21} and ξ_{22} and$\cdots$and ξ_{2m} then η_{21} and η_{22} and$\cdots$and η_{2n}
> $\cdots$
> If ξ_{k1} and ξ_{k2} and$\cdots$and ξ_{km} then η_{k1} and η_{k2} and$\cdots$and η_{kn}

$$\tag{8.37}$$

and the inference rule

$$
\eta_j^* = \sum_{i=1}^{k} \frac{c_i \cdot \eta_{ij}\big|_{(\alpha_1 \rhd \xi_{i1}) \cap (\alpha_2 \rhd \xi_{i2}) \cap \cdots \cap (\alpha_m \rhd \xi_{im})}}{c_1 + c_2 + \cdots + c_k} \tag{8.38}
$$

for $j = 1, 2, \cdots, n$, where the coefficients are determined by

$$
c_i = \mathcal{M}\left\{(\alpha_1 \rhd \xi_{i1}) \cap (\alpha_2 \rhd \xi_{i2}) \cap \cdots \cap (\alpha_m \rhd \xi_{im})\right\} \tag{8.39}
$$

for $i = 1, 2, \cdots, k$. Thus we obtain

$$
\beta_j = E[\eta_j^*] \tag{8.40}
$$

for $j = 1, 2, \cdots, n$. Until now we have constructed a function from inputs $\alpha_1, \alpha_2, \cdots, \alpha_m$ to outputs $\beta_1, \beta_2, \cdots, \beta_n$. Write this function by f, i.e.,

$$
(\beta_1, \beta_2, \cdots, \beta_n) = f(\alpha_1, \alpha_2, \cdots, \alpha_m). \tag{8.41}
$$

Then we get an uncertain system f.

Theorem 8.6 *Assume that $\xi_{i1}, \xi_{i2}, \cdots, \xi_{im}, \eta_{i1}, \eta_{i2}, \cdots, \eta_{in}$ are independent uncertain sets with continuous first identification functions $\lambda_{i1}, \lambda_{i2}, \cdots, \lambda_{im},$*
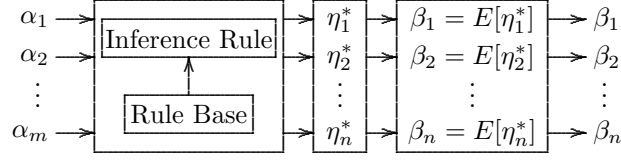
Figure 8.4: An Uncertain System

$\nu_{i1}, \nu_{i2}, \cdots, \nu_{in}$, $i = 1, 2, \cdots, k$, respectively. Then the uncertain system from $(\alpha_1, \alpha_2, \cdots, \alpha_m)$ to $(\beta_1, \beta_2, \cdots, \beta_n)$ is

$$\beta_j = \sum_{i=1}^{k} \frac{c_i \cdot E[\eta_{ij}^*]}{c_1 + c_2 + \cdots + c_k} \tag{8.42}$$

for $j = 1, 2, \cdots, n$, where c_i are constants determined by

$$c_i = \begin{cases} \min_{1 \le l \le m} \lambda_{il}(\alpha_l), & if \; \min_{1 \le l \le m} \lambda_{il}(\alpha_l) < 0.5 \\ 1, & if \; \min_{1 \le l \le m} \lambda_{il}(\alpha_l) = 0.5 \end{cases} \tag{8.43}$$

and η_{ij}^* are uncertain sets whose first identification functions are given by

$$\nu_{ij}^*(z) = \begin{cases} \dfrac{\nu_{ij}(z)}{\min_{1 \le l \le m} \lambda_{il}(\alpha_l)} \wedge 0.5, & if \; \min_{1 \le l \le m} \lambda_{il}(\alpha_l) < 0.5 \\ \nu_{ij}(z), & if \; \min_{1 \le l \le m} \lambda_{il}(\alpha_l) = 0.5 \end{cases} \tag{8.44}$$

for $i = 1, 2, \cdots, k$, $j = 1, 2, \cdots, n$, respectively.

Proof: It follows from inference rule 8.4 that the uncertain sets η_j^* are

$$\eta_j^* = \sum_{i=1}^{k} \frac{c_i \cdot \eta_{ij}^*}{c_1 + c_2 + \cdots + c_k}$$

for $j = 1, 2, \cdots, n$. Since $\eta_{ij}^*, i = 1, 2, \cdots, k, j = 1, 2, \cdots, n$ are independent uncertain sets, we get the theorem immediately by the linearity of expected value operator.

Remark 8.1: The uncertain system allows the uncertain sets η_{ij} in the rule-base (8.37) become constants b_{ij}, i.e.,

$$\eta_{ij} = b_{ij} \tag{8.45}$$

for $i = 1, 2, \cdots, k$ and $j = 1, 2, \cdots, n$. For this case, the uncertain system (8.42) becomes

$$\beta_j = \sum_{i=1}^{k} \frac{c_i \cdot b_{ij}}{c_1 + c_2 + \cdots + c_k} \tag{8.46}$$

for $j = 1, 2, \cdots, n$.

Remark 8.2: The uncertain system allows the uncertain sets η_{ij} in the rule-base (8.37) become functions h_{ij} of inputs $\alpha_1, \alpha_2, \cdots, \alpha_m$, i.e.,

$$\eta_{ij} = h_{ij}(\alpha_1, \alpha_2, \cdots, \alpha_m) \qquad (8.47)$$

for $i = 1, 2, \cdots, k$ and $j = 1, 2, \cdots, n$. For this case, the uncertain system (8.42) becomes

$$\beta_j = \sum_{i=1}^{k} \frac{c_i \cdot h_{ij}(\alpha_1, \alpha_2, \cdots, \alpha_m)}{c_1 + c_2 + \cdots + c_k} \qquad (8.48)$$

for $j = 1, 2, \cdots, n$.

Uncertain Systems are Universal Approximators

Theorem 8.7 *For any given continuous function g on a compact set $D \subset \Re^m$ and any given $\varepsilon > 0$, there exists an uncertain system f such that*

$$\sup_{(\alpha_1, \alpha_2, \cdots, \alpha_m) \in D} \| f(\alpha_1, \alpha_2, \cdots, \alpha_m) - g(\alpha_1, \alpha_2, \cdots, \alpha_m) \| < \varepsilon. \qquad (8.49)$$

8.8 Uncertain Inference Control

By an *uncertain inference controller* we mean a controller based on uncertain inference rule. Figure 8.5 shows an uncertain inference control system consisting of an uncertain inference controller and a process. Note that t represents time, $\alpha_1(t), \alpha_2(t), \cdots, \alpha_m(t)$ are not only the inputs of uncertain inference controller but also the outputs of process, and $\beta_1(t), \beta_2(t), \cdots, \beta_n(t)$ are not only the outputs of uncertain inference controller but also the inputs of process.
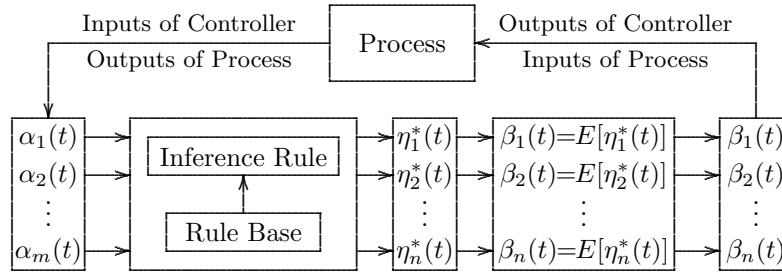
Figure 8.5: An Uncertain Inference Control System

8.9 Inverted Pendulum

Inverted pendulum system is a nonlinear unstable system that is widely used as a benchmark for testing control algorithms. Many good techniques already exist for balancing inverted pendulum. It is possible to use an uncertain inference controller to balance an inverted pendulum.

8.10 Autonomous Vehicle

8.11 Temperature Control

Chapter 9

Supplements

This chapter introduces the generalized law of excluded middle and the maximum uncertainty principle that have been used in this book. This chapter also provides a brief history of evolution of measures and discusses why uncertainty theory is more reasonable than fuzzy set theory. Finally, we answer the question "what is uncertainty".

9.1 Generalized Law of Excluded Middle

The law of excluded middle tells us that a proposition is either true or false, and the law of contradiction tells us that a proposition cannot be both true and false.

In the state of uncertainty, some people said, the law of excluded middle is no longer valid because the truth degree of a proposition is no longer 0 or 1. I cannot gainsay this viewpoint to a certain extent. But it does not mean that you might "go as you please". I think at least the generalized law of excluded middle should be valid in the state of uncertainty. That is, the truth degree should meet

$$\mathcal{M}\{\Lambda\} + \mathcal{M}\{\Lambda^c\} = 1$$

for each proposition Λ. In other words, the generalized law of excluded middle is nothing but the self-duality property. Although the generalized law of excluded middle is not just the classical one, both of them agree with each other when the uncertainty vanishes, i.e., when $\mathcal{M}\{\Lambda\}$ tends to 0 or 1.

9.2 Maximum Uncertainty Principle

An event has no uncertainty if its measure is 1 (or 0) because we may believe that the event occurs (or not). An event is the most uncertain if its measure

is 0.5 because the event and its complement may be regarded as "equally likely".

In practice, if there is no information about the measure of an event, we should assign 0.5 to it. Sometimes, only partial information is available. For this case, the value of measure may be specified in some range. What value does the measure take? For the safety purpose, we should assign it the value as close to 0.5 as possible. This is the maximum uncertainty principle proposed by Liu [127].

Maximum Uncertainty Principle: *For any event, if there are multiple reasonable values that an uncertain measure may take, then the value as close to 0.5 as possible is assigned to the event.*

Example 9.1: Let Λ be an event. Based on some given information, the measure value $\mathcal{M}\{\Lambda\}$ is on the interval $[a, b]$. By using the maximum uncertainty principle, we should assign

$$
\mathcal{M}\{\Lambda\} = \left\{
\begin{array}{ll}
a, & \text{if } 0.5 < a \le b \\
0.5, & \text{if } a \le 0.5 \le b \\
b, & \text{if } a \le b < 0.5.
\end{array}
\right.
$$

9.3   Evolution of Measures

Let us begin with the technical term "event" in uncertainty theory, or equivalently, "proposition" in logic. The following are examples of event,

"It will rain tomorrow";
"Tianjin is not far from Beijing".

Perhaps we are accustomed to assigning to each event two state symbols, true or false, to represent the truth degree of the event. However, starting with Aristotle's epoch, it was observed that such a two-valued assignment has its shortages. Today, it is well-known that the truth degree may be assigned any values between 0 and 1, where 0 represents "completely false" and 1 represents "completely true". The higher the truth degree is, the more true the event is.

It is clear that there are multiple assignment ways. This fact has resulted in several types of measure, for example,

1933: Probability Measure (A.N. Kolmogoroff);
1954: Capacity (G. Choquet);
1974: Fuzzy Measure (M. Sugeno);
1978: Possibility Measure (L.A. Zadeh);
2002: Credibility Measure (B. Liu and Y. Liu);
2007: Uncertain Measure (B. Liu).

Probability Measure

A classical measure is essentially a set function satisfying nonnegativity and countable additivity axioms. In order to deal with randomness, Kolmogoroff (1933) defined a probability measure as a special classical measure with normality axiom. That is, the following three axioms must be satisfied:

Axiom 1. *(Normality)* $\Pr\{\Omega\} = 1$ *for the universal set* Ω.

Axiom 2. *(Nonnegativity)* $\Pr\{A\} \geq 0$ *for any event A.*

Axiom 3. *(Countable Additivity) For every countable sequence of mutually disjoint events* $\{A_i\}$*, we have*

$$\Pr\left\{\bigcup_{i=1}^{\infty} A_i\right\} = \sum_{i=1}^{\infty} \Pr\{A_i\}. \tag{9.1}$$

A lot of surveys showed that probability measure is successful in physical systems but not applicable to human systems.

Capacity

In order to deal with human systems, the additivity axiom seems too strong. The earliest challenge was from the theory of capacities by Choquet (1954) in which only monotonicity and continuity axioms were assumed. However, capacity is inconsistent with the law of excluded middle.

Fuzzy Measure

Sugeno (1974) generalized classical measure theory to fuzzy measure theory by replacing additivity axiom with weaker axioms of monotonicity and semi-continuity. However, fuzzy measure is inconsistent with the law of excluded middle.

Possibility Measure

In order to deal with fuzziness, Zadeh (1978) proposed a possibility measure that satisfies the following axioms:

Axiom 1. *(Normality)* $\mathrm{Pos}\{\Theta\} = 1$ *for the universal set* Θ.

Axiom 1. *(Nonnegativity)* $\mathrm{Pos}\{\emptyset\} = 0$ *for the empty set* $\emptyset$.

Axiom 3. *(Maximality) For every sequence of events* $\{A_i\}$*, we have*

$$\mathrm{Pos}\left\{\bigcup_{i=1}^{\infty} A_i\right\} = \bigvee_{i=1}^{\infty} \mathrm{Pos}\{A_i\}. \tag{9.2}$$

Unfortunately, possibility measure is inconsistent with the law of excluded middle.

Credibility Measure

In order to overcome the shortage of possibility measure, Liu and Liu (2002) presented a credibility measure that may be defined by the following four axioms:

Axiom 1. *(Normality)* $\mathrm{Cr}\{\Theta\} = 1$ *for the universal set* Θ.

Axiom 2. *(Monotonicity)* $\mathrm{Cr}\{A\} \leq \mathrm{Cr}\{B\}$ *whenever* $A \subset B$.

Axiom 3. *(Self-Duality)* $\mathrm{Cr}\{A\} + \mathrm{Cr}\{A^c\} = 1$ *for any event* A.

Axiom 4. *(Maximality) For every sequence of events* $\{A_i\}$*, we have*

$$\mathrm{Cr}\left\{\bigcup_{i=1}^{\infty} A_i\right\} = \bigvee_{i=1}^{\infty} \mathrm{Cr}\{A_i\}, \quad if \bigvee_{i=1}^{\infty} \mathrm{Cr}\{A_i\} < 0.5. \qquad (9.3)$$

Although credibility measure assumes the self-duality axiom and is consistent with the law of excluded middle, it seems that there is no phenomena in the real world behaving fuzziness. This issue will be discussed in next section.

Uncertain Measure

In order to deal with uncertainty in human systems, Liu (2007) presented an uncertain measure based on the following four axioms:

Axiom 1. *(Normality)* $\mathcal{M}\{\Gamma\} = 1$ *for the universal set* Γ.

Axiom 2. *(Monotonicity)* $\mathcal{M}\{\Lambda_1\} \leq \mathcal{M}\{\Lambda_2\}$ *whenever* $\Lambda_1 \subset \Lambda_2$.

Axiom 3. *(Self-Duality)* $\mathcal{M}\{\Lambda\} + \mathcal{M}\{\Lambda^c\} = 1$ *for any event* Λ.

Axiom 4. *(Countable Subadditivity) For every countable sequence of events* $\{\Lambda_i\}$*, we have*

$$\mathcal{M}\left\{\bigcup_{i=1}^{\infty} \Lambda_i\right\} \leq \sum_{i=1}^{\infty} \mathcal{M}\{\Lambda_i\}. \qquad (9.4)$$

Uncertain measure is neither a completely additive measure nor a completely nonadditive measure. In fact, uncertain measure is a "partially additive measure" because of its self-duality. One advantage of uncertain measure is the well consistency with the law of excluded middle.

9.4   Uncertainty or Fuzziness?

The essential differentia between fuzziness and uncertainty is that the former assumes

$$\mathrm{Pos}\{A \cup B\} = \mathrm{Pos}\{A\} \vee \mathrm{Pos}\{B\}$$

or equivalently,

$$\mathrm{Cr}\{A \cup B\} = \mathrm{Cr}\{A\} \vee \mathrm{Cr}\{B\}, \quad \text{if } \mathrm{Cr}\{A\} \vee \mathrm{Cr}\{B\} < 0.5$$

for any events A and B no matter if they are independent or not, and the latter assumes

$$\mathcal{M}\{A \cup B\} = \mathcal{M}\{A\} \vee \mathcal{M}\{B\}$$

only for independent events A and B. However, a lot of surveys showed that the measure of union of events is not necessarily the maximum measure, i.e.,

$$\mathcal{M}\{A \cup B\} \neq \mathcal{M}\{A\} \vee \mathcal{M}\{B\}$$

when the events A and B are not independent. This fact states that human systems do not behave fuzziness and the tool of membership function is no longer helpful. Fuzziness is an extremely special but non-practical case of uncertainty. This is the main reason why we need the uncertainty theory.

9.5   What is Uncertainty?

Now we are arriving at the end of this book. Perhaps some readers may complain that I never clarify what uncertainty is. In fact, I really have no idea how to use natural language to define uncertainty clearly, and I think all existing definitions by natural language are specious. A very personal and ultra viewpoint is that the words like randomness, fuzziness, roughness, vagueness, greyness, and uncertainty are nothing but ambiguity of human language!

However, fortunately, some "mathematical scales" have been invented to measure the truth degree of an event, for example, probability measure, capacity, fuzzy measure, possibility measure, credibility measure as well as uncertain measure. All of those measures may be defined clearly and precisely by axiomatic methods.

Let us go back to the first question "what is uncertainty". Perhaps we can answer it this way. If it happened that some phenomena can be quantified by uncertain measure, then we call the phenomena "uncertainty". In other words, uncertainty is any concept that satisfies the axioms of uncertainty theory. Thus there are various valid possibilities (*e.g.*, a personal belief degree) to interpret uncertainty theory. Could you agree with me? I hope that uncertainty theory may play a mathematical model of uncertainty in your own problem.

Appendix A

Probability Theory

Probability theory is a branch of mathematics for studying the behavior of random phenomena. The emphasis in this appendix is mainly on probability space, random variable, probability distribution, independence, expected value, variance, moments, critical values, entropy and conditional probability. The main results in this appendix are well-known. For this reason the credit references are not provided.

A.1 Probability Space

Let Ω be a nonempty set, and $\mathcal{A}$ a σ-algebra over Ω. If Ω is countable, usually $\mathcal{A}$ is the power set of Ω. If Ω is uncountable, for example $\Omega = [0, 1]$, usually $\mathcal{A}$ is the Borel algebra of Ω. Each element in $\mathcal{A}$ is called an event. In order to present an axiomatic definition of probability, it is necessary to assign to each event A a number $\Pr\{A\}$ which indicates the probability that A will occur. In order to ensure that the number $\Pr\{A\}$ has certain mathematical properties which we intuitively expect a probability to have, the following three axioms must be satisfied:

Axiom 1. *(Normality)* $\Pr\{\Omega\} = 1$.

Axiom 2. *(Nonnegativity)* $\Pr\{A\} \geq 0$ *for any event A.*

Axiom 3. *(Countable Additivity) For every countable sequence of mutually disjoint events $\{A_i\}$, we have*

$$\Pr\left\{\bigcup_{i=1}^{\infty} A_i\right\} = \sum_{i=1}^{\infty} \Pr\{A_i\}. \tag{A.1}$$

Definition A.1 *The set function $\Pr$ is called a probability measure if it satisfies the normality, nonnegativity, and countable additivity axioms.*

Example A.1: Let $\Omega = \{\omega_1, \omega_2, \cdots\}$, and let $\mathcal{A}$ be the power set of Ω. Assume that $p_1, p_2, \cdots$ are nonnegative numbers such that $p_1 + p_2 + \cdots = 1$. Define a set function on $\mathcal{A}$ as

$$\Pr\{A\} = \sum_{\omega_i \in A} p_i, \quad A \in \mathcal{A}. \tag{A.2}$$

Then Pr is a probability measure.

Example A.2: Let ϕ be a nonnegative and integrable function on $\Re$ (the set of real numbers) such that

$$\int_{\Re} \phi(x)\mathrm{d}x = 1. \tag{A.3}$$

Then for any Borel set A, the set function

$$\Pr\{A\} = \int_A \phi(x)\mathrm{d}x \tag{A.4}$$

is a probability measure on $\Re$.

Theorem A.1 *Let Ω be a nonempty set, $\mathcal{A}$ a σ-algebra over Ω, and $\Pr$ a probability measure. Then we have*
(a) $\Pr$ is self-dual, i.e., $\Pr\{A\} + \Pr\{A^c\} = 1$ for any $A \in \mathcal{A}$;
(b) $\Pr$ is increasing, i.e., $\Pr\{A\} \leq \Pr\{B\}$ whenever $A \subset B$.

Proof: (a) Since A and A^c are disjoint events and $A \cup A^c = \Omega$, we have $\Pr\{A\} + \Pr\{A^c\} = \Pr\{\Omega\} = 1$. (b) Since $A \subset B$, we have $B = A \cup (B \cap A^c)$, where A and $B \cap A^c$ are disjoint events. Therefore $\Pr\{B\} = \Pr\{A\} + \Pr\{B \cap A^c\} \geq \Pr\{A\}$.

Probability Continuity Theorem

Theorem A.2 *(Probability Continuity Theorem) Let Ω be a nonempty set, $\mathcal{A}$ a σ-algebra over Ω, and $\Pr$ a probability measure. If $A_1, A_2, \cdots \in \mathcal{A}$ and $\lim_{i \to \infty} A_i$ exists, then*

$$\lim_{i \to \infty} \Pr\{A_i\} = \Pr\left\{\lim_{i \to \infty} A_i\right\}. \tag{A.5}$$

Proof: STEP 1: Suppose $\{A_i\}$ is an increasing sequence. Write $A_i \to A$ and $A_0 = \emptyset$. Then $\{A_i \backslash A_{i-1}\}$ is a sequence of disjoint events and

$$\bigcup_{i=1}^{\infty}(A_i \backslash A_{i-1}) = A, \quad \bigcup_{i=1}^{k}(A_i \backslash A_{i-1}) = A_k$$

for $k = 1, 2, \cdots$ Thus we have

$$
\Pr\{A\} = \Pr\left\{\bigcup_{i=1}^{\infty}(A_i \backslash A_{i-1})\right\} = \sum_{i=1}^{\infty} \Pr\{A_i \backslash A_{i-1}\}
$$

$$
= \lim_{k \to \infty} \sum_{i=1}^{k} \Pr\{A_i \backslash A_{i-1}\} = \lim_{k \to \infty} \Pr\left\{\bigcup_{i=1}^{k}(A_i \backslash A_{i-1})\right\}
$$

$$
= \lim_{k \to \infty} \Pr\{A_k\}.
$$

STEP 2: If $\{A_i\}$ is a decreasing sequence, then the sequence $\{A_1 \backslash A_i\}$ is clearly increasing. It follows that

$$
\Pr\{A_1\} - \Pr\{A\} = \Pr\left\{\lim_{i \to \infty}(A_1 \backslash A_i)\right\} = \lim_{i \to \infty} \Pr\{A_1 \backslash A_i\}
$$

$$
= \Pr\{A_1\} - \lim_{i \to \infty} \Pr\{A_i\}
$$

which implies that $\Pr\{A_i\} \to \Pr\{A\}$.

STEP 3: If $\{A_i\}$ is a sequence of events such that $A_i \to A$, then for each k, we have

$$
\bigcap_{i=k}^{\infty} A_i \subset A_k \subset \bigcup_{i=k}^{\infty} A_i.
$$

Since Pr is increasing, we have

$$
\Pr\left\{\bigcap_{i=k}^{\infty} A_i\right\} \leq \Pr\{A_k\} \leq \Pr\left\{\bigcup_{i=k}^{\infty} A_i\right\}.
$$

Note that

$$
\bigcap_{i=k}^{\infty} A_i \uparrow A, \quad \bigcup_{i=k}^{\infty} A_i \downarrow A.
$$

It follows from Steps 1 and 2 that $\Pr\{A_i\} \to \Pr\{A\}$.

Probability Space

Definition A.2 *Let Ω be a nonempty set, $\mathcal{A}$ a σ-algebra over Ω, and Pr a probability measure. Then the triplet $(\Omega, \mathcal{A}, \Pr)$ is called a probability space.*

Example A.3: Let $\Omega = \{\omega_1, \omega_2, \cdots\}$, $\mathcal{A}$ the power set of Ω, and Pr a probability measure defined by (A.2). Then $(\Omega, \mathcal{A}, \Pr)$ is a probability space.

Example A.4: Let $\Omega = [0, 1]$, $\mathcal{A}$ the Borel algebra over Ω, and Pr the Lebesgue measure. Then $([0, 1], \mathcal{A}, \Pr)$ is a probability space, and sometimes is called *Lebesgue unit interval*. For many purposes it is sufficient to use it as the basic probability space.

Product Probability Space

Let $(\Omega_i, \mathcal{A}_i, \mathrm{Pr}_i)$, $i = 1, 2, \cdots, n$ be probability spaces, and $\Omega = \Omega_1 \times \Omega_2 \times \cdots \times \Omega_n$, $\mathcal{A} = \mathcal{A}_1 \times \mathcal{A}_2 \times \cdots \times \mathcal{A}_n$. Note that the probability measures $\mathrm{Pr}_i, i = 1, 2, \cdots, n$ are finite. It follows from the classical measure theory that there is a unique measure Pr on $\mathcal{A}$ such that

$$\mathrm{Pr}\{A_1 \times A_2 \times \cdots \times A_n\} = \mathrm{Pr}_1\{A_1\} \times \mathrm{Pr}_2\{A_2\} \times \cdots \times \mathrm{Pr}_n\{A_n\}$$

for any $A_i \in \mathcal{A}_i$, $i = 1, 2, \cdots, n$. This conclusion is called the *product probability theorem*. The measure Pr is also a probability measure since

$$\mathrm{Pr}\{\Omega\} = \mathrm{Pr}_1\{\Omega_1\} \times \mathrm{Pr}_2\{\Omega_2\} \times \cdots \times \mathrm{Pr}_n\{\Omega_n\} = 1.$$

Such a probability measure is called the product probability measure, denoted by $\mathrm{Pr} = \mathrm{Pr}_1 \times \mathrm{Pr}_2 \times \cdots \times \mathrm{Pr}_n$.

Definition A.3 *Let $(\Omega_i, \mathcal{A}_i, \mathrm{Pr}_i)$, $i = 1, 2, \cdots, n$ be probability spaces, and $\Omega = \Omega_1 \times \Omega_2 \times \cdots \times \Omega_n$, $\mathcal{A} = \mathcal{A}_1 \times \mathcal{A}_2 \times \cdots \times \mathcal{A}_n$, $\mathrm{Pr} = \mathrm{Pr}_1 \times \mathrm{Pr}_2 \times \cdots \times \mathrm{Pr}_n$. Then the triplet $(\Omega, \mathcal{A}, \mathrm{Pr})$ is called the product probability space.*

A.2   Random Variable

Definition A.4 *A random variable is a measurable function from a probability space $(\Omega, \mathcal{A}, \mathrm{Pr})$ to the set of real numbers, i.e., for any Borel set B of real numbers, the set*

$$\{\xi \in B\} = \{\omega \in \Omega \mid \xi(\omega) \in B\} \tag{A.6}$$

is an event.

Example A.5: Take $(\Omega, \mathcal{A}, \mathrm{Pr})$ to be $\{\omega_1, \omega_2\}$ with $\mathrm{Pr}\{\omega_1\} = \mathrm{Pr}\{\omega_2\} = 0.5$. Then the function

$$\xi(\omega) = \begin{cases} 0, & \text{if } \omega = \omega_1 \\ 1, & \text{if } \omega = \omega_2 \end{cases}$$

is a random variable.

Example A.6: Take $(\Omega, \mathcal{A}, \mathrm{Pr})$ to be the interval $[0, 1]$ with Borel algebra and Lebesgue measure. We define ξ as an identity function from Ω to $[0,1]$. Since ξ is a measurable function, it is a random variable.

Example A.7: A deterministic number c may be regarded as a special random variable. In fact, it is the constant function $\xi(\omega) \equiv c$ on the probability space $(\Omega, \mathcal{A}, \mathrm{Pr})$.

Definition A.5 *Let ξ_1 and ξ_2 be random variables defined on the probability space $(\Omega, \mathcal{A}, \mathrm{Pr})$. We say $\xi_1 = \xi_2$ if $\xi_1(\omega) = \xi_2(\omega)$ for almost all $\omega \in \Omega$.*

Random Vector

Definition A.6 *An n-dimensional random vector is a measurable function from a probability space* $(\Omega, \mathcal{A}, \mathrm{Pr})$ *to the set of n-dimensional real vectors, i.e., for any Borel set B of* $\Re^n$, *the set*

$$\{\boldsymbol{\xi} \in B\} = \big\{\omega \in \Omega \mid \boldsymbol{\xi}(\omega) \in B\big\} \tag{A.7}$$

is an event.

Theorem A.3 *The vector* $(\xi_1, \xi_2, \cdots, \xi_n)$ *is a random vector if and only if* $\xi_1, \xi_2, \cdots, \xi_n$ *are random variables.*

Proof: Write $\boldsymbol{\xi} = (\xi_1, \xi_2, \cdots, \xi_n)$. Suppose that $\boldsymbol{\xi}$ is a random vector on the probability space $(\Omega, \mathcal{A}, \mathrm{Pr})$. For any Borel set B of $\Re$, the set $B \times \Re^{n-1}$ is also a Borel set of $\Re^n$. Thus we have

$$\{\xi_1 \in B\} = \{\xi_1 \in B, \xi_2 \in \Re, \cdots, \xi_n \in \Re\} = \{\boldsymbol{\xi} \in B \times \Re^{n-1}\} \in \mathcal{A}$$

which implies that ξ_1 is a random variable. A similar process may prove that $\xi_2, \xi_3, \cdots, \xi_n$ are random variables. Conversely, suppose that all $\xi_1, \xi_2, \cdots, \xi_n$ are random variables on the probability space $(\Omega, \mathcal{A}, \mathrm{Pr})$. We define

$$\mathcal{B} = \big\{ B \subset \Re^n \mid \{\boldsymbol{\xi} \in B\} \in \mathcal{A} \big\}.$$

The vector $\boldsymbol{\xi} = (\xi_1, \xi_2, \cdots, \xi_n)$ is proved to be a random vector if we can prove that $\mathcal{B}$ contains all Borel sets of $\Re^n$. First, the class $\mathcal{B}$ contains all open intervals of $\Re^n$ because

$$\left\{\boldsymbol{\xi} \in \prod_{i=1}^{n}(a_i, b_i)\right\} = \bigcap_{i=1}^{n} \{\xi_i \in (a_i, b_i)\} \in \mathcal{A}.$$

Next, the class $\mathcal{B}$ is a σ-algebra of $\Re^n$ because (i) we have $\Re^n \in \mathcal{B}$ since $\{\boldsymbol{\xi} \in \Re^n\} = \Omega \in \mathcal{A}$; (ii) if $B \in \mathcal{B}$, then $\{\boldsymbol{\xi} \in B\} \in \mathcal{A}$, and

$$\{\boldsymbol{\xi} \in B^c\} = \{\boldsymbol{\xi} \in B\}^c \in \mathcal{A}$$

which implies that $B^c \in \mathcal{B}$; (iii) if $B_i \in \mathcal{B}$ for $i = 1, 2, \cdots$, then $\{\boldsymbol{\xi} \in B_i\} \in \mathcal{A}$ and

$$\left\{\boldsymbol{\xi} \in \bigcup_{i=1}^{\infty} B_i\right\} = \bigcup_{i=1}^{\infty} \{\boldsymbol{\xi} \in B_i\} \in \mathcal{A}$$

which implies that $\cup_i B_i \in \mathcal{B}$. Since the smallest σ-algebra containing all open intervals of $\Re^n$ is just the Borel algebra of $\Re^n$, the class $\mathcal{B}$ contains all Borel sets of $\Re^n$. The theorem is proved.

Random Arithmetic

In this subsections, we will suppose that all random variables are defined on a common probability space. Otherwise, we may embed them into the product probability space.

Definition A.7 *Let $f : \Re^n \to \Re$ be a measurable function, and $\xi_1, \xi_2, \cdots, \xi_n$ random variables defined on the probability space $(\Omega, \mathcal{A}, \mathrm{Pr})$. Then $\xi = f(\xi_1, \xi_2, \cdots, \xi_n)$ is a random variable defined by*

$$\xi(\omega) = f(\xi_1(\omega), \xi_2(\omega), \cdots, \xi_n(\omega)), \quad \forall \omega \in \Omega. \tag{A.8}$$

Example A.8: Let ξ_1 and ξ_2 be random variables on the probability space $(\Omega, \mathcal{A}, \mathrm{Pr})$. Then their sum is

$$(\xi_1 + \xi_2)(\omega) = \xi_1(\omega) + \xi_2(\omega), \quad \forall \omega \in \Omega$$

and their product is

$$(\xi_1 \times \xi_2)(\omega) = \xi_1(\omega) \times \xi_2(\omega), \quad \forall \omega \in \Omega.$$

The reader may wonder whether $\xi(\omega_1, \omega_2, \cdots, \omega_n)$ defined by (A.7) is a random variable. The following theorem answers this question.

Theorem A.4 *Let $\boldsymbol{\xi}$ be an n-dimensional random vector, and $f : \Re^n \to \Re$ a measurable function. Then $f(\boldsymbol{\xi})$ is a random variable.*

Proof: Assume that $\boldsymbol{\xi}$ is a random vector on the probability space $(\Omega, \mathcal{A}, \mathrm{Pr})$. For any Borel set B of $\Re$, since f is a measurable function, $f^{-1}(B)$ is also a Borel set of $\Re^n$. Thus we have

$$\{f(\boldsymbol{\xi}) \in B\} = \left\{\boldsymbol{\xi} \in f^{-1}(B)\right\} \in \mathcal{A}$$

which implies that $f(\boldsymbol{\xi})$ is a random variable.

A.3   Probability Distribution

Definition A.8 *The probability distribution $\Phi: \Re \to [0,1]$ of a random variable ξ is defined by*

$$\Phi(x) = \mathrm{Pr}\left\{\xi \leq x\right\}. \tag{A.9}$$

That is, $\Phi(x)$ is the probability that the random variable ξ takes a value less than or equal to x.

Example A.9: Take $(\Omega, \mathcal{A}, \mathrm{Pr})$ to be $\{\omega_1, \omega_2\}$ with $\mathrm{Pr}\{\omega_1\} = \mathrm{Pr}\{\omega_2\} = 0.5$. We now define a random variable as follows,

$$\xi(\omega) = \begin{cases} -1, & \text{if } \omega = \omega_1 \\ 1, & \text{if } \omega = \omega_2. \end{cases}$$

Then ξ has a probability distribution

$$\Phi(x) = \begin{cases} 0, & \text{if } x < -1 \\ 0.5, & \text{if } -1 \leq x < 1 \\ 1, & \text{if } x \geq 1. \end{cases}$$

Theorem A.5 *(Sufficient and Necessary Condition for Probability Distribution) A function $\Phi : \Re \to [0,1]$ is a probability distribution if and only if it is an increasing and right-continuous function with*

$$\lim_{x \to -\infty} \Phi(x) = 0; \qquad \lim_{x \to +\infty} \Phi(x) = 1. \tag{A.10}$$

Proof: For any $x, y \in \Re$ with $x < y$, we have

$$\Phi(y) - \Phi(x) = \Pr\{x < \xi \leq y\} \geq 0.$$

Thus the probability distribution Φ is increasing. Next, let $\{\varepsilon_i\}$ be a sequence of positive numbers such that $\varepsilon_i \to 0$ as $i \to \infty$. Then, for every $i \geq 1$, we have

$$\Phi(x + \varepsilon_i) - \Phi(x) = \Pr\{x < \xi \leq x + \varepsilon_i\}.$$

It follows from the probability continuity theorem that

$$\lim_{i \to \infty} \Phi(x + \varepsilon_i) - \Phi(x) = \Pr\{\emptyset\} = 0.$$

Hence Φ is a right-continuous function. Finally,

$$\lim_{x \to -\infty} \Phi(x) = \lim_{x \to -\infty} \Pr\{\xi \leq x\} = \Pr\{\emptyset\} = 0,$$

$$\lim_{x \to +\infty} \Phi(x) = \lim_{x \to +\infty} \Pr\{\xi \leq x\} = \Pr\{\Omega\} = 1.$$

Conversely, it is known there is a unique probability measure Pr on the Borel algebra over $\Re$ such that $\Pr\{(-\infty, x]\} = \Phi(x)$ for all $x \in \Re$. Furthermore, it is easy to verify that the random variable defined by $\xi(x) = x$ from the probability space $(\Re, \mathcal{A}, \Pr)$ to $\Re$ has the probability distribution Φ.

Probability Density Function

Definition A.9 *The probability density function $\phi : \Re \to [0, +\infty)$ of a random variable ξ is a function such that*

$$\Phi(x) = \int_{-\infty}^{x} \phi(y) \mathrm{d}y \tag{A.11}$$

holds for all $x \in \Re$, where Φ is the probability distribution of the random variable ξ.

Theorem A.6 *(Probability Inversion Theorem) Let ξ be a random variable whose probability density function ϕ exists. Then for any Borel set B of $\Re$, we have*

$$\Pr\{\xi \in B\} = \int_B \phi(y)\mathrm{d}y. \tag{A.12}$$

Proof: Let $\mathcal{C}$ be the class of all subsets C of $\Re$ for which the relation

$$\Pr\{\xi \in C\} = \int_C \phi(y)\mathrm{d}y \tag{A.13}$$

holds. We will show that $\mathcal{C}$ contains all Borel sets of $\Re$. It follows from the probability continuity theorem and relation (A.13) that $\mathcal{C}$ is a monotone class. It is also clear that $\mathcal{C}$ contains all intervals of the form $(-\infty, a]$, $(a, b]$, (b, ∞) and $\emptyset$ since

$$\Pr\{\xi \in (-\infty, a]\} = \Phi(a) = \int_{-\infty}^{a} \phi(y)\mathrm{d}y,$$

$$\Pr\{\xi \in (b, +\infty)\} = \Phi(+\infty) - \Phi(b) = \int_{b}^{+\infty} \phi(y)\mathrm{d}y,$$

$$\Pr\{\xi \in (a, b]\} = \Phi(b) - \Phi(a) = \int_{a}^{b} \phi(y)\mathrm{d}y,$$

$$\Pr\{\xi \in \emptyset\} = 0 = \int_{\emptyset} \phi(y)\mathrm{d}y$$

where Φ is the probability distribution of ξ. Let $\mathcal{F}$ be the algebra consisting of all finite unions of disjoint sets of the form $(-\infty, a]$, $(a, b]$, (b, ∞) and $\emptyset$. Note that for any disjoint sets $C_1, C_2, \cdots, C_m$ of $\mathcal{F}$ and $C = C_1 \cup C_2 \cup \cdots \cup C_m$, we have

$$\Pr\{\xi \in C\} = \sum_{j=1}^{m} \Pr\{\xi \in C_j\} = \sum_{j=1}^{m} \int_{C_j} \phi(y)\mathrm{d}y = \int_{C} \phi(y)\mathrm{d}y.$$

That is, $C \in \mathcal{C}$. Hence we have $\mathcal{F} \subset \mathcal{C}$. Since the smallest σ-algebra containing $\mathcal{F}$ is just the Borel algebra of $\Re$, the monotone class theorem implies that $\mathcal{C}$ contains all Borel sets of $\Re$.

Some Special Distributions

Uniform Distribution: A random variable ξ has a uniform distribution if its probability density function is defined by

$$\phi(x) = \frac{1}{b-a}, \quad a \le x \le b. \tag{A.14}$$

where a and b are given real numbers with $a < b$.

Exponential Distribution: A random variable ξ has an exponential distribution if its probability density function is defined by

$$\phi(x) = \frac{1}{\beta} \exp\left(-\frac{x}{\beta}\right), \quad x \geq 0 \tag{A.15}$$

where β is a positive number.

Normal Distribution: A random variable ξ has a normal distribution if its probability density function is defined by

$$\phi(x) = \frac{1}{\sigma\sqrt{2\pi}} \exp\left(-\frac{(x-\mu)^2}{2\sigma^2}\right), \quad x \in \Re \tag{A.16}$$

where μ and σ are real numbers.

A.4 Independence

Definition A.10 *The random variables $\xi_1, \xi_2, \cdots, \xi_m$ are said to be independent if*

$$\Pr\left\{\bigcap_{i=1}^{m}\{\xi_i \in B_i\}\right\} = \prod_{i=1}^{m} \Pr\{\xi_i \in B_i\} \tag{A.17}$$

for any Borel sets $B_1, B_2, \cdots, B_m$ of $\Re$.

Theorem A.7 *Let ξ_i be random variables with probability distributions Φ_i, $i = 1, 2, \cdots, m$, respectively, and Φ the probability distribution of the random vector $(\xi_1, \xi_2, \cdots, \xi_m)$. Then $\xi_1, \xi_2, \cdots, \xi_m$ are independent if and only if*

$$\Phi(x_1, x_2, \cdots, x_m) = \Phi_1(x_1)\Phi_2(x_2)\cdots\Phi_m(x_m) \tag{A.18}$$

for all $(x_1, x_2, \cdots, x_m) \in \Re^m$.

Proof: If $\xi_1, \xi_2, \cdots, \xi_m$ are independent random variables, then we have

$$\begin{aligned}
\Phi(x_1, x_2, \cdots, x_m) &= \Pr\{\xi_1 \leq x_1, \xi_2 \leq x_2, \cdots, \xi_m \leq x_m\} \\
&= \Pr\{\xi_1 \leq x_1\}\Pr\{\xi_2 \leq x_2\}\cdots\Pr\{\xi_m \leq x_m\} \\
&= \Phi_1(x_1)\Phi_2(x_2)\cdots\Phi_m(x_m)
\end{aligned}$$

for all $(x_1, x_2, \cdots, x_m) \in \Re^m$. Conversely, assume that (A.18) holds. Let $x_2, x_3, \cdots, x_m$ be fixed real numbers, and $\mathcal{C}$ the class of all subsets C of $\Re$ for which the relation

$$\Pr\{\xi_1 \in C, \xi_2 \leq x_2, \cdots, \xi_m \leq x_m\} = \Pr\{\xi_1 \in C\}\prod_{i=2}^{m}\Pr\{\xi_i \leq x_i\} \tag{A.19}$$

holds. We will show that $\mathcal{C}$ contains all Borel sets of $\Re$. It follows from the probability continuity theorem and relation (A.19) that $\mathcal{C}$ is a monotone

class. It is also clear that $\mathcal{C}$ contains all intervals of the form $(-\infty, a]$, $(a, b]$, (b, ∞) and $\emptyset$. Let $\mathcal{F}$ be the algebra consisting of all finite unions of disjoint sets of the form $(-\infty, a]$, $(a, b]$, (b, ∞) and $\emptyset$. Note that for any disjoint sets $C_1, C_2, \cdots, C_k$ of $\mathcal{F}$ and $C = C_1 \cup C_2 \cup \cdots \cup C_k$, we have

$$\Pr\{\xi_1 \in C, \xi_2 \le x_2, \cdots, \xi_m \le x_m\}$$

$$= \sum_{j=1}^{m} \Pr\{\xi_1 \in C_j, \xi_2 \le x_2, \cdots, \xi_m \le x_m\}$$

$$= \Pr\{\xi_1 \in C\} \Pr\{\xi_2 \le x_2\} \cdots \Pr\{\xi_m \le x_m\}.$$

That is, $C \in \mathcal{C}$. Hence we have $\mathcal{F} \subset \mathcal{C}$. Since the smallest σ-algebra containing $\mathcal{F}$ is just the Borel algebra of $\Re$, the monotone class theorem implies that $\mathcal{C}$ contains all Borel sets of $\Re$. Applying the same reasoning to each ξ_i in turn, we obtain the independence of the random variables.

Theorem A.8 *Let ξ_i be random variables with probability density functions ϕ_i, $i = 1, 2, \cdots, m$, respectively, and ϕ the probability density function of the random vector $(\xi_1, \xi_2, \cdots, \xi_m)$. Then $\xi_1, \xi_2, \cdots, \xi_m$ are independent if and only if*

$$\phi(x_1, x_2, \cdots, x_m) = \phi_1(x_1)\phi_2(x_2) \cdots \phi_m(x_m) \qquad (A.20)$$

for almost all $(x_1, x_2, \cdots, x_m) \in \Re^m$.

Proof: If $\phi(x_1, x_2, \cdots, x_m) = \phi_1(x_1)\phi_2(x_2) \cdots \phi_m(x_m)$ a.e., then we have

$$\begin{aligned}
\Phi(x_1, x_2, \cdots, x_m) &= \int_{-\infty}^{x_1} \int_{-\infty}^{x_2} \cdots \int_{-\infty}^{x_m} \phi(t_1, t_2, \cdots, t_m) \mathrm{d}t_1 \mathrm{d}t_2 \cdots \mathrm{d}t_m \\
&= \int_{-\infty}^{x_1} \int_{-\infty}^{x_2} \cdots \int_{-\infty}^{x_m} \phi_1(t_1)\phi_2(t_2) \cdots \phi_m(t_m) \mathrm{d}t_1 \mathrm{d}t_2 \cdots \mathrm{d}t_m \\
&= \int_{-\infty}^{x_1} \phi_1(t_1) \mathrm{d}t_1 \int_{-\infty}^{x_2} \phi_2(t_2) \mathrm{d}t_2 \cdots \int_{-\infty}^{x_m} \phi_m(t_m) \mathrm{d}t_m \\
&= \Phi_1(x_1)\Phi_2(x_2) \cdots \Phi_m(x_m)
\end{aligned}$$

for all $(x_1, x_2, \cdots, x_m) \in \Re^m$. Thus $\xi_1, \xi_2, \cdots, \xi_m$ are independent. Conversely, if $\xi_1, \xi_2, \cdots, \xi_m$ are independent, then for any $(x_1, x_2, \cdots, x_m) \in \Re^m$, we have $\Phi(x_1, x_2, \cdots, x_m) = \Phi_1(x_1)\Phi_2(x_2) \cdots \Phi_m(x_m)$. Hence

$$\Phi(x_1, x_2, \cdots, x_m) = \int_{-\infty}^{x_1} \int_{-\infty}^{x_2} \cdots \int_{-\infty}^{x_m} \phi_1(t_1)\phi_2(t_2) \cdots \phi_m(t_m) \mathrm{d}t_1 \mathrm{d}t_2 \cdots \mathrm{d}t_m$$

which implies that $\phi(x_1, x_2, \cdots, x_m) = \phi_1(x_1)\phi_2(x_2) \cdots \phi_m(x_m)$ a.e.

Example A.10: Let $\xi_1, \xi_2, \cdots, \xi_m$ be independent random variables with probability density functions $\phi_1, \phi_2, \cdots, \phi_m$, respectively, and $f : \Re^m \to \Re$ a measurable function. Then for any Borel set B of real numbers, the probability $\Pr\{f(\xi_1, \xi_2, \cdots, \xi_m) \in B\}$ is

$$\iint \cdots \int_{f(x_1, x_2, \cdots, x_m) \in B} \phi_1(x_1)\phi_2(x_2) \cdots \phi_m(x_m) \mathrm{d}x_1 \mathrm{d}x_2 \cdots \mathrm{d}x_m.$$

A.5 Expected Value

Definition A.11 *Let ξ be a random variable. Then the expected value of ξ is defined by*

$$E[\xi] = \int_0^{+\infty} \Pr\{\xi \geq r\}\mathrm{d}r - \int_{-\infty}^0 \Pr\{\xi \leq r\}\mathrm{d}r \qquad (A.21)$$

provided that at least one of the two integrals is finite.

Example A.11: Assume that ξ is a discrete random variable taking values x_i with probabilities p_i, $i = 1, 2, \cdots, m$, respectively. It follows from the definition of expected value operator that

$$E[\xi] = \sum_{i=1}^m p_i x_i.$$

Theorem A.9 *Let ξ be a random variable whose probability density function ϕ exists. If the Lebesgue integral*

$$\int_{-\infty}^{+\infty} x\phi(x)\mathrm{d}x$$

is finite, then we have

$$E[\xi] = \int_{-\infty}^{+\infty} x\phi(x)\mathrm{d}x. \qquad (A.22)$$

Proof: It follows from Definition A.11 and Fubini Theorem that

$$
\begin{aligned}
E[\xi] &= \int_0^{+\infty} \Pr\{\xi \geq r\}\mathrm{d}r - \int_{-\infty}^0 \Pr\{\xi \leq r\}\mathrm{d}r \\
&= \int_0^{+\infty} \left[\int_r^{+\infty} \phi(x)\mathrm{d}x\right]\mathrm{d}r - \int_{-\infty}^0 \left[\int_{-\infty}^r \phi(x)\mathrm{d}x\right]\mathrm{d}r \\
&= \int_0^{+\infty} \left[\int_0^x \phi(x)\mathrm{d}r\right]\mathrm{d}x - \int_{-\infty}^0 \left[\int_x^0 \phi(x)\mathrm{d}r\right]\mathrm{d}x \\
&= \int_0^{+\infty} x\phi(x)\mathrm{d}x + \int_{-\infty}^0 x\phi(x)\mathrm{d}x \\
&= \int_{-\infty}^{+\infty} x\phi(x)\mathrm{d}x.
\end{aligned}
$$

The theorem is proved.

Theorem A.10 *Let ξ be a random variable with probability distribution Φ. If the Lebesgue-Stieltjes integral*

$$\int_{-\infty}^{+\infty} x \mathrm{d}\Phi(x)$$

is finite, then we have

$$E[\xi] = \int_{-\infty}^{+\infty} x \mathrm{d}\Phi(x). \tag{A.23}$$

Proof: Since the Lebesgue-Stieltjes integral $\int_{-\infty}^{+\infty} x \mathrm{d}\Phi(x)$ is finite, we immediately have

$$\lim_{y\to+\infty} \int_0^y x \mathrm{d}\Phi(x) = \int_0^{+\infty} x \mathrm{d}\Phi(x), \quad \lim_{y\to-\infty} \int_y^0 x \mathrm{d}\Phi(x) = \int_{-\infty}^0 x \mathrm{d}\Phi(x)$$

and

$$\lim_{y\to+\infty} \int_y^{+\infty} x \mathrm{d}\Phi(x) = 0, \quad \lim_{y\to-\infty} \int_{-\infty}^y x \mathrm{d}\Phi(x) = 0.$$

It follows from

$$\int_y^{+\infty} x \mathrm{d}\Phi(x) \geq y \left(\lim_{z\to+\infty} \Phi(z) - \Phi(y) \right) = y(1 - \Phi(y)) \geq 0, \quad \text{if } y > 0,$$

$$\int_{-\infty}^y x \mathrm{d}\Phi(x) \leq y \left(\Phi(y) - \lim_{z\to-\infty} \Phi(z) \right) = y\Phi(y) \leq 0, \quad \text{if } y < 0$$

that

$$\lim_{y\to+\infty} y \left(1 - \Phi(y) \right) = 0, \quad \lim_{y\to-\infty} y\Phi(y) = 0.$$

Let $0 = x_0 < x_1 < x_2 < \cdots < x_n = y$ be a partition of $[0, y]$. Then we have

$$\sum_{i=0}^{n-1} x_i \left(\Phi(x_{i+1}) - \Phi(x_i) \right) \to \int_0^y x \mathrm{d}\Phi(x)$$

and

$$\sum_{i=0}^{n-1} (1 - \Phi(x_{i+1}))(x_{i+1} - x_i) \to \int_0^y \Pr\{\xi \geq r\} \mathrm{d}r$$

as $\max\{|x_{i+1} - x_i| : i = 0, 1, \cdots, n - 1\} \to 0$. Since

$$\sum_{i=0}^{n-1} x_i \left(\Phi(x_{i+1}) - \Phi(x_i) \right) - \sum_{i=0}^{n-1} (1 - \Phi(x_{i+1}))(x_{i+1} - x_i) = y(\Phi(y) - 1) \to 0$$

as $y \to +\infty$. This fact implies that

$$\int_0^{+\infty} \Pr\{\xi \geq r\} \mathrm{d}r = \int_0^{+\infty} x \mathrm{d}\Phi(x).$$

A similar way may prove that

$$-\int_{-\infty}^{0} \Pr\{\xi \leq r\} dr = \int_{-\infty}^{0} x d\Phi(x).$$

Thus (A.23) is verified by the above two equations.

Linearity of Expected Value Operator

Theorem A.11 *Let ξ and η be random variables with finite expected values. Then for any numbers a and b, we have*

$$E[a\xi + b\eta] = aE[\xi] + bE[\eta]. \tag{A.24}$$

Proof: STEP 1: We first prove that $E[\xi + b] = E[\xi] + b$ for any real number b. When $b \geq 0$, we have

$$
\begin{aligned}
E[\xi + b] &= \int_{0}^{\infty} \Pr\{\xi + b \geq r\} dr - \int_{-\infty}^{0} \Pr\{\xi + b \leq r\} dr \\
&= \int_{0}^{\infty} \Pr\{\xi \geq r - b\} dr - \int_{-\infty}^{0} \Pr\{\xi \leq r - b\} dr \\
&= E[\xi] + \int_{0}^{b} \left(\Pr\{\xi \geq r - b\} + \Pr\{\xi < r - b\} \right) dr \\
&= E[\xi] + b.
\end{aligned}
$$

If $b < 0$, then we have

$$E[\xi + b] = E[\xi] - \int_{b}^{0} \left(\Pr\{\xi \geq r - b\} + \Pr\{\xi < r - b\} \right) dr = E[\xi] + b.$$

STEP 2: We prove that $E[a\xi] = aE[\xi]$ for any real number a. If $a = 0$, then the equation $E[a\xi] = aE[\xi]$ holds trivially. If $a > 0$, we have

$$
\begin{aligned}
E[a\xi] &= \int_{0}^{\infty} \Pr\{a\xi \geq r\} dr - \int_{-\infty}^{0} \Pr\{a\xi \leq r\} dr \\
&= \int_{0}^{\infty} \Pr\left\{\xi \geq \frac{r}{a}\right\} dr - \int_{-\infty}^{0} \Pr\left\{\xi \leq \frac{r}{a}\right\} dr \\
&= a \int_{0}^{\infty} \Pr\left\{\xi \geq \frac{r}{a}\right\} d\left(\frac{r}{a}\right) - a \int_{-\infty}^{0} \Pr\left\{\xi \leq \frac{r}{a}\right\} d\left(\frac{r}{a}\right) \\
&= aE[\xi].
\end{aligned}
$$

If $a < 0$, we have

$$
\begin{aligned}
E[a\xi] &= \int_0^\infty \Pr\{a\xi \geq r\}\mathrm{d}r - \int_{-\infty}^0 \Pr\{a\xi \leq r\}\mathrm{d}r \\
&= \int_0^\infty \Pr\left\{\xi \leq \frac{r}{a}\right\}\mathrm{d}r - \int_{-\infty}^0 \Pr\left\{\xi \geq \frac{r}{a}\right\}\mathrm{d}r \\
&= a\int_0^\infty \Pr\left\{\xi \geq \frac{r}{a}\right\}\mathrm{d}\left(\frac{r}{a}\right) - a\int_{-\infty}^0 \Pr\left\{\xi \leq \frac{r}{a}\right\}\mathrm{d}\left(\frac{r}{a}\right) \\
&= aE[\xi].
\end{aligned}
$$

STEP 3: We prove that $E[\xi + \eta] = E[\xi] + E[\eta]$ when both ξ and η are nonnegative simple random variables taking values $a_1, a_2, \cdots, a_m$ and $b_1, b_2, \cdots, b_n$, respectively. Then $\xi + \eta$ is also a nonnegative simple random variable taking values $a_i + b_j$, $i = 1, 2, \cdots, m$, $j = 1, 2, \cdots, n$. Thus we have

$$
\begin{aligned}
E[\xi + \eta] &= \sum_{i=1}^m \sum_{j=1}^n (a_i + b_j)\Pr\{\xi = a_i, \eta = b_j\} \\
&= \sum_{i=1}^m \sum_{j=1}^n a_i \Pr\{\xi = a_i, \eta = b_j\} + \sum_{i=1}^m \sum_{j=1}^n b_j \Pr\{\xi = a_i, \eta = b_j\} \\
&= \sum_{i=1}^m a_i \Pr\{\xi = a_i\} + \sum_{j=1}^n b_j \Pr\{\eta = b_j\} \\
&= E[\xi] + E[\eta].
\end{aligned}
$$

STEP 4: We prove that $E[\xi + \eta] = E[\xi] + E[\eta]$ when both ξ and η are nonnegative random variables. For every $i \geq 1$ and every $\omega \in \Omega$, we define

$$
\xi_i(\omega) = \begin{cases} \dfrac{k-1}{2^i}, & \text{if } \dfrac{k-1}{2^i} \leq \xi(\omega) < \dfrac{k}{2^i},\ k = 1, 2, \cdots, i2^i \\[2ex] i, & \text{if } i \leq \xi(\omega), \end{cases}
$$

$$
\eta_i(\omega) = \begin{cases} \dfrac{k-1}{2^i}, & \text{if } \dfrac{k-1}{2^i} \leq \eta(\omega) < \dfrac{k}{2^i},\ k = 1, 2, \cdots, i2^i \\[2ex] i, & \text{if } i \leq \eta(\omega). \end{cases}
$$

Then $\{\xi_i\}$, $\{\eta_i\}$ and $\{\xi_i + \eta_i\}$ are three sequences of nonnegative simple random variables such that $\xi_i \uparrow \xi$, $\eta_i \uparrow \eta$ and $\xi_i + \eta_i \uparrow \xi + \eta$ as $i \to \infty$. Note that the functions $\Pr\{\xi_i > r\}$, $\Pr\{\eta_i > r\}$, $\Pr\{\xi_i + \eta_i > r\}$, $i = 1, 2, \cdots$ are also simple. It follows from the probability continuity theorem that

$$
\Pr\{\xi_i > r\} \uparrow \Pr\{\xi > r\}, \forall r \geq 0
$$

as $i \to \infty$. Since the expected value $E[\xi]$ exists, we have

$$
E[\xi_i] = \int_0^{+\infty} \Pr\{\xi_i > r\}\mathrm{d}r \to \int_0^{+\infty} \Pr\{\xi > r\}\mathrm{d}r = E[\xi]
$$

as $i \to \infty$. Similarly, we may prove that $E[\eta_i] \to E[\eta]$ and $E[\xi_i + \eta_i] \to E[\xi + \eta]$ as $i \to \infty$. It follows from Step 3 that $E[\xi + \eta] = E[\xi] + E[\eta]$.

STEP 5: We prove that $E[\xi + \eta] = E[\xi] + E[\eta]$ when ξ and η are arbitrary random variables. Define

$$\xi_i(\omega) = \begin{cases} \xi(\omega), & \text{if } \xi(\omega) \geq -i \\ -i, & \text{otherwise,} \end{cases} \qquad \eta_i(\omega) = \begin{cases} \eta(\omega), & \text{if } \eta(\omega) \geq -i \\ -i, & \text{otherwise.} \end{cases}$$

Since the expected values $E[\xi]$ and $E[\eta]$ are finite, we have

$$\lim_{i \to \infty} E[\xi_i] = E[\xi], \quad \lim_{i \to \infty} E[\eta_i] = E[\eta], \quad \lim_{i \to \infty} E[\xi_i + \eta_i] = E[\xi + \eta].$$

Note that $(\xi_i + i)$ and $(\eta_i + i)$ are nonnegative random variables. It follows from Steps 1 and 4 that

$$
\begin{aligned}
E[\xi + \eta] &= \lim_{i \to \infty} E[\xi_i + \eta_i] \\
&= \lim_{i \to \infty} \left(E[(\xi_i + i) + (\eta_i + i)] - 2i \right) \\
&= \lim_{i \to \infty} \left(E[\xi_i + i] + E[\eta_i + i] - 2i \right) \\
&= \lim_{i \to \infty} \left(E[\xi_i] + i + E[\eta_i] + i - 2i \right) \\
&= \lim_{i \to \infty} E[\xi_i] + \lim_{i \to \infty} E[\eta_i] \\
&= E[\xi] + E[\eta].
\end{aligned}
$$

STEP 6: The linearity $E[a\xi + b\eta] = aE[\xi] + bE[\eta]$ follows immediately from Steps 2 and 5. The theorem is proved.

Product of Independent Random Variables

Theorem A.12 *Let ξ and η be independent random variables with finite expected values. Then the expected value of $\xi\eta$ exists and*

$$E[\xi\eta] = E[\xi]E[\eta]. \tag{A.25}$$

Proof: STEP 1: We first prove the case where both ξ and η are nonnegative simple random variables taking values $a_1, a_2, \cdots, a_m$ and $b_1, b_2, \cdots, b_n$, respectively. Then $\xi\eta$ is also a nonnegative simple random variable taking values $a_i b_j$, $i = 1, 2, \cdots, m$, $j = 1, 2, \cdots, n$. It follows from the independence of ξ and η that

$$
\begin{aligned}
E[\xi\eta] &= \sum_{i=1}^{m} \sum_{j=1}^{n} a_i b_j \Pr\{\xi = a_i, \eta = b_j\} \\
&= \sum_{i=1}^{m} \sum_{j=1}^{n} a_i b_j \Pr\{\xi = a_i\} \Pr\{\eta = b_j\} \\
&= \left(\sum_{i=1}^{m} a_i \Pr\{\xi = a_i\} \right) \left(\sum_{j=1}^{n} b_j \Pr\{\eta = b_j\} \right) \\
&= E[\xi]E[\eta].
\end{aligned}
$$

STEP 2: Next we prove the case where ξ and η are nonnegative random variables. For every $i \geq 1$ and every $\omega \in \Omega$, we define

$$\xi_i(\omega) = \begin{cases} \dfrac{k-1}{2^i}, & \text{if } \dfrac{k-1}{2^i} \leq \xi(\omega) < \dfrac{k}{2^i}, \ k = 1, 2, \cdots, i2^i \\[2ex] i, & \text{if } i \leq \xi(\omega), \end{cases}$$

$$\eta_i(\omega) = \begin{cases} \dfrac{k-1}{2^i}, & \text{if } \dfrac{k-1}{2^i} \leq \eta(\omega) < \dfrac{k}{2^i}, \ k = 1, 2, \cdots, i2^i \\[2ex] i, & \text{if } i \leq \eta(\omega). \end{cases}$$

Then $\{\xi_i\}$, $\{\eta_i\}$ and $\{\xi_i\eta_i\}$ are three sequences of nonnegative simple random variables such that $\xi_i \uparrow \xi$, $\eta_i \uparrow \eta$ and $\xi_i\eta_i \uparrow \xi\eta$ as $i \to \infty$. It follows from the independence of ξ and η that ξ_i and η_i are independent. Hence we have $E[\xi_i\eta_i] = E[\xi_i]E[\eta_i]$ for $i = 1, 2, \cdots$ It follows from the probability continuity theorem that $\Pr\{\xi_i > r\}, i = 1, 2, \cdots$ are simple functions such that

$$\Pr\{\xi_i > r\} \uparrow \Pr\{\xi > r\}, \ \text{ for all } r \geq 0$$

as $i \to \infty$. Since the expected value $E[\xi]$ exists, we have

$$E[\xi_i] = \int_0^{+\infty} \Pr\{\xi_i > r\}\mathrm{d}r \to \int_0^{+\infty} \Pr\{\xi > r\}\mathrm{d}r = E[\xi]$$

as $i \to \infty$. Similarly, we may prove that $E[\eta_i] \to E[\eta]$ and $E[\xi_i\eta_i] \to E[\xi\eta]$ as $i \to \infty$. Therefore $E[\xi\eta] = E[\xi]E[\eta]$.

STEP 3: Finally, if ξ and η are arbitrary independent random variables, then the nonnegative random variables ξ^+ and η^+ are independent and so are ξ^+ and η^-, ξ^- and η^+, ξ^- and η^-. Thus we have

$$E[\xi^+\eta^+] = E[\xi^+]E[\eta^+], \quad E[\xi^+\eta^-] = E[\xi^+]E[\eta^-],$$
$$E[\xi^-\eta^+] = E[\xi^-]E[\eta^+], \quad E[\xi^-\eta^-] = E[\xi^-]E[\eta^-].$$

It follows that

$$\begin{aligned} E[\xi\eta] &= E[(\xi^+ - \xi^-)(\eta^+ - \eta^-)] \\ &= E[\xi^+\eta^+] - E[\xi^+\eta^-] - E[\xi^-\eta^+] + E[\xi^-\eta^-] \\ &= E[\xi^+]E[\eta^+] - E[\xi^+]E[\eta^-] - E[\xi^-]E[\eta^+] + E[\xi^-]E[\eta^-] \\ &= \left(E[\xi^+] - E[\xi^-]\right)\left(E[\eta^+] - E[\eta^-]\right) \\ &= E[\xi^+ - \xi^-]E[\eta^+ - \eta^-] \\ &= E[\xi]E[\eta] \end{aligned}$$

which proves the theorem.

A.6   Variance

Definition A.12 *Let ξ be a random variable with finite expected value e. Then the variance of ξ is defined by $V[\xi] = E[(\xi - e)^2]$.*

The variance of a random variable provides a measure of the spread of the distribution around its expected value. A small value of variance indicates that the random variable is tightly concentrated around its expected value; and a large value of variance indicates that the random variable has a wide spread around its expected value.

Theorem A.13 *If ξ is a random variable whose variance exists, a and b are real numbers, then $V[a\xi + b] = a^2 V[\xi]$.*

Proof: It follows from the definition of variance that

$$V[a\xi + b] = E\left[(a\xi + b - aE[\xi] - b)^2\right] = a^2 E[(\xi - E[\xi])^2] = a^2 V[\xi].$$

Theorem A.14 *Let ξ be a random variable with expected value e. Then $V[\xi] = 0$ if and only if $\Pr\{\xi = e\} = 1$.*

Proof: If $V[\xi] = 0$, then $E[(\xi - e)^2] = 0$. Thus we have

$$\int_0^{+\infty} \Pr\{(\xi - e)^2 \geq r\}\mathrm{d}r = 0$$

which implies $\Pr\{(\xi-e)^2 \geq r\} = 0$ for any $r > 0$. Hence we have $\Pr\{(\xi-e)^2 = 0\} = 1$, i.e., $\Pr\{\xi = e\} = 1$. Conversely, if $\Pr\{\xi = e\} = 1$, then we have $\Pr\{(\xi - e)^2 = 0\} = 1$ and $\Pr\{(\xi - e)^2 \geq r\} = 0$ for any $r > 0$. Thus

$$V[\xi] = \int_0^{+\infty} \Pr\{(\xi - e)^2 \geq r\}\mathrm{d}r = 0.$$

Theorem A.15 *If $\xi_1, \xi_2, \cdots, \xi_n$ are independent random variables with finite expected values, then*

$$V[\xi_1 + \xi_2 + \cdots + \xi_n] = V[\xi_1] + V[\xi_2] + \cdots + V[\xi_n]. \qquad (A.26)$$

Proof: It follows from the definition of variance that

$$V\left[\sum_{i=1}^n \xi_i\right] = E\left[(\xi_1 + \xi_2 + \cdots + \xi_n - E[\xi_1] - E[\xi_2] - \cdots - E[\xi_n])^2\right]$$

$$= \sum_{i=1}^n E\left[(\xi_i - E[\xi_i])^2\right] + 2\sum_{i=1}^{n-1}\sum_{j=i+1}^n E\left[(\xi_i - E[\xi_i])(\xi_j - E[\xi_j])\right].$$

Since $\xi_1, \xi_2, \cdots, \xi_n$ are independent, $E\left[(\xi_i - E[\xi_i])(\xi_j - E[\xi_j])\right] = 0$ for all i, j with $i \neq j$. Thus (A.26) holds.

A.7   Moments

Definition A.13 *Let ξ be a random variable, and k a positive number. Then*
(a) the expected value $E[\xi^k]$ is called the kth moment;
(b) the expected value $E[|\xi|^k]$ is called the kth absolute moment;
(c) the expected value $E[(\xi - E[\xi])^k]$ is called the kth central moment;
(d) the expected value $E[|\xi - E[\xi]|^k]$ is called the kth absolute central moment.

Note that the first central moment is always 0, the first moment is just the expected value, and the second central moment is just the variance.

Theorem A.16 *Let ξ be a nonnegative random variable, and k a positive number. Then the k-th moment*

$$E[\xi^k] = k \int_0^{+\infty} r^{k-1} \Pr\{\xi \geq r\} \mathrm{d}r. \qquad (A.27)$$

Proof: It follows from the nonnegativity of ξ that

$$E[\xi^k] = \int_0^\infty \Pr\{\xi^k \geq x\} \mathrm{d}x = \int_0^\infty \Pr\{\xi \geq r\} \mathrm{d}r^k = k \int_0^\infty r^{k-1} \Pr\{\xi \geq r\} \mathrm{d}r.$$

The theorem is proved.

A.8   Critical Values

Let ξ be a random variable. In order to measure it, we may use its expected value. Alternately, we may employ α-optimistic value and α-pessimistic value as a ranking measure.

Definition A.14 *Let ξ be a random variable, and $\alpha \in (0, 1]$. Then*

$$\xi_{\sup}(\alpha) = \sup \left\{ r \mid \Pr\{\xi \geq r\} \geq \alpha \right\} \qquad (A.28)$$

is called the α-optimistic value of ξ, and

$$\xi_{\inf}(\alpha) = \inf \left\{ r \mid \Pr\{\xi \leq r\} \geq \alpha \right\} \qquad (A.29)$$

is called the α-pessimistic value of ξ.

This means that the random variable ξ will reach upwards of the α-optimistic value $\xi_{\sup}(\alpha)$ at least α of time, and will be below the α-pessimistic value $\xi_{\inf}(\alpha)$ at least α of time. The optimistic value is also called percentile.

Theorem A.17 *Let ξ be a random variable, and $\alpha \in (0, 1]$. Then we have*

$$\Pr\{\xi \geq \xi_{\sup}(\alpha)\} \geq \alpha, \quad \Pr\{\xi \leq \xi_{\inf}(\alpha)\} \geq \alpha \qquad (A.30)$$

where $\xi_{\sup}(\alpha)$ and $\xi_{\inf}(\alpha)$ are the α-optimistic and α-pessimistic values of the random variable ξ, respectively.

Proof: It follows from the definition of the optimistic value that there exists an increasing sequence $\{r_i\}$ such that $\Pr\{\xi \geq r_i\} \geq \alpha$ and $r_i \uparrow \xi_{\sup}(\alpha)$ as $i \to \infty$. Since $\{\xi \geq r_i\} \downarrow \{\xi \geq \xi_{\sup}(\alpha)\}$, it follows from the probability continuity theorem that

$$\Pr\{\xi \geq \xi_{\sup}(\alpha)\} = \lim_{i \to \infty} \Pr\{\xi \geq r_i\} \geq \alpha.$$

The inequality $\Pr\{\xi \leq \xi_{\inf}(\alpha)\} \geq \alpha$ may be proved similarly.

Theorem A.18 *Let ξ be a random variable, and $\alpha \in (0,1]$. Then we have*
(a) $\xi_{\inf}(\alpha)$ is an increasing and left-continuous function of α;
(b) $\xi_{\sup}(\alpha)$ is a decreasing and left-continuous function of α.

Proof: (a) Let α_1 and α_2 be two numbers with $0 < \alpha_1 < \alpha_2 \leq 1$. Then for any number $r < \xi_{\sup}(\alpha_2)$, we have $\Pr\{\xi \geq r\} \geq \alpha_2 > \alpha_1$. Thus, by the definition of optimistic value, we obtain $\xi_{\sup}(\alpha_1) \geq \xi_{\sup}(\alpha_2)$. That is, the value $\xi_{\sup}(\alpha)$ is a decreasing function of α. Next, we prove the left-continuity of $\xi_{\inf}(\alpha)$ with respect to α. Let $\{\alpha_i\}$ be an arbitrary sequence of positive numbers such that $\alpha_i \uparrow \alpha$. Then $\{\xi_{\inf}(\alpha_i)\}$ is an increasing sequence. If the limitation is equal to $\xi_{\inf}(\alpha)$, then the left-continuity is proved. Otherwise, there exists a number z^* such that

$$\lim_{i \to \infty} \xi_{\inf}(\alpha_i) < z^* < \xi_{\inf}(\alpha).$$

Thus $\Pr\{\xi \leq z^*\} \geq \alpha_i$ for each i. Letting $i \to \infty$, we get $\Pr\{\xi \leq z^*\} \geq \alpha$. Hence $z^* \geq \xi_{\inf}(\alpha)$. A contradiction proves the left-continuity of $\xi_{\inf}(\alpha)$ with respect to α. The part (b) may be proved similarly.

A.9   Entropy

Given a random variable, what is the degree of difficulty of predicting the specified value that the random variable will take? In order to answer this question, Shannon [199] defined a concept of entropy as a measure of uncertainty.

Entropy of Discrete Random Variables

Definition A.15 *Let ξ be a discrete random variable taking values x_i with probabilities p_i, $i = 1, 2, \cdots$, respectively. Then its entropy is defined by*

$$H[\xi] = -\sum_{i=1}^{\infty} p_i \ln p_i. \tag{A.31}$$

It should be noticed that the entropy depends only on the number of values and their probabilities and does not depend on the actual values that the random variable takes.
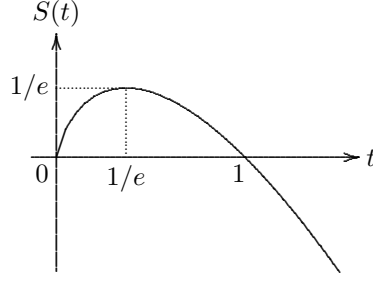
Figure A.1: Function $S(t) = -t \ln t$ is concave

Theorem A.19 *Let ξ be a discrete random variable taking values x_i with probabilities p_i, $i = 1, 2, \cdots$, respectively. Then*

$$H[\xi] \geq 0 \qquad\qquad (A.32)$$

and equality holds if and only if there exists an index k such that $p_k = 1$, i.e., ξ is essentially a deterministic number.

Proof: The nonnegativity is clear. In addition, $H[\xi] = 0$ if and only if $p_i = 0$ or 1 for each i. That is, there exists one and only one index k such that $p_k = 1$. The theorem is proved.

This theorem states that the entropy of a discrete random variable reaches its minimum 0 when the random variable degenerates to a deterministic number. In this case, there is no uncertainty.

Theorem A.20 *Let ξ be a simple random variable taking values x_i with probabilities p_i, $i = 1, 2, \cdots, n$, respectively. Then*

$$H[\xi] \leq \ln n \qquad\qquad (A.33)$$

and equality holds if and only if $p_i \equiv 1/n$ for all $i = 1, 2, \cdots, n$.

Proof: Since the function $S(t)$ is a concave function of t and $p_1 + p_2 + \cdots + p_n = 1$, we have

$$-\sum_{i=1}^{n} p_i \ln p_i \leq -n \left(\frac{1}{n} \sum_{i=1}^{n} p_i \right) \ln \left(\frac{1}{n} \sum_{i=1}^{n} p_i \right) = \ln n$$

which implies that $H[\xi] \leq \ln n$ and equality holds if and only if $p_1 = p_2 = \cdots = p_n$, i.e., $p_i \equiv 1/n$ for all $i = 1, 2, \cdots, n$.

This theorem states that the entropy of a simple random variable reaches its maximum $\ln n$ when all outcomes are equiprobable. In this case, there is no preference among all the values that the random variable will take.

Entropy of Absolutely Continuous Random Variables

Definition A.16 *Let ξ be a random variable with probability density function ϕ. Then its entropy is defined by*

$$H[\xi] = - \int_{-\infty}^{+\infty} \phi(x) \ln \phi(x) \mathrm{d}x. \tag{A.34}$$

Example A.12: Let ξ be a uniformly distributed random variable on $[a, b]$. Then its entropy is $H[\xi] = \ln(b - a)$. This example shows that the entropy of absolutely continuous random variable may assume both positive and negative values since $\ln(b - a) < 0$ if $b - a < 1$; and $\ln(b - a) > 0$ if $b - a > 1$.

Example A.13: Let ξ be an exponentially distributed random variable with expected value β. Then its entropy is $H[\xi] = 1 + \ln \beta$.

Example A.14: Let ξ be a normally distributed random variable with expected value e and variance σ^2. Then its entropy is $H[\xi] = 1/2 + \ln \sqrt{2\pi}\sigma$.

Maximum Entropy Principle

Given some constraints, for example, expected value and variance, there are usually multiple compatible probability distributions. For this case, we would like to select the distribution that maximizes the value of entropy and satisfies the prescribed constraints. This method is often referred to as the *maximum entropy principle* (Jaynes [64]).

Example A.15: Let ξ be an absolutely continuous random variable on $[a, b]$. The maximum entropy principle attempts to find the probability density function $\phi(x)$ that maximizes the entropy

$$- \int_a^b \phi(x) \ln \phi(x) \mathrm{d}x$$

subject to the natural constraint $\int_a^b \phi(x) \mathrm{d}x = 1$. The Lagrangian is

$$L = - \int_a^b \phi(x) \ln \phi(x) \mathrm{d}x - \lambda \left(\int_a^b \phi(x) \mathrm{d}x - 1 \right).$$

It follows from the Euler-Lagrange equation that the maximum entropy probability density function meets

$$\ln \phi(x) + 1 + \lambda = 0$$

and has the form $\phi(x) = \exp(-1 - \lambda)$. Substituting it into the natural constraint, we get

$$\phi^*(x) = \frac{1}{b - a}, \quad a \le x \le b$$

which is just the uniformly distributed random variable, and the maximum entropy is $H[\xi^*] = \ln(b - a)$.

Example A.16: Let ξ be an absolutely continuous random variable on $[0, \infty)$. Assume that the expected value of ξ is prescribed to be β. The maximum entropy probability density function $\phi(x)$ should maximize the entropy

$$-\int_0^{+\infty} \phi(x) \ln \phi(x) \mathrm{d}x$$

subject to the constraints

$$\int_0^{+\infty} \phi(x) \mathrm{d}x = 1, \quad \int_0^{+\infty} x\phi(x) \mathrm{d}x = \beta.$$

The Lagrangian is

$$L = -\int_0^{\infty} \phi(x) \ln \phi(x) \mathrm{d}x - \lambda_1 \left(\int_0^{\infty} \phi(x) \mathrm{d}x - 1 \right) - \lambda_2 \left(\int_0^{\infty} x\phi(x) \mathrm{d}x - \beta \right).$$

The maximum entropy probability density function meets Euler-Lagrange equation

$$\ln \phi(x) + 1 + \lambda_1 + \lambda_2 x = 0$$

and has the form $\phi(x) = \exp(-1 - \lambda_1 - \lambda_2 x)$. Substituting it into the constraints, we get

$$\phi^*(x) = \frac{1}{\beta} \exp\left(-\frac{x}{\beta} \right), \quad x \geq 0$$

which is just the exponentially distributed random variable, and the maximum entropy is $H[\xi^*] = 1 + \ln \beta$.

Example A.17: Let ξ be an absolutely continuous random variable on $(-\infty, +\infty)$. Assume that the expected value and variance of ξ are prescribed to be μ and σ^2, respectively. The maximum entropy probability density function $\phi(x)$ should maximize the entropy

$$-\int_{-\infty}^{+\infty} \phi(x) \ln \phi(x) \mathrm{d}x$$

subject to the constraints

$$\int_{-\infty}^{+\infty} \phi(x) \mathrm{d}x = 1, \quad \int_{-\infty}^{+\infty} x\phi(x) \mathrm{d}x = \mu, \quad \int_{-\infty}^{+\infty} (x - \mu)^2 \phi(x) \mathrm{d}x = \sigma^2.$$

The Lagrangian is

$$L = -\int_{-\infty}^{+\infty} \phi(x) \ln \phi(x) \mathrm{d}x - \lambda_1 \left(\int_{-\infty}^{+\infty} \phi(x) \mathrm{d}x - 1 \right)$$

$$-\lambda_2 \left(\int_{-\infty}^{+\infty} x\phi(x) \mathrm{d}x - \mu \right) - \lambda_3 \left(\int_{-\infty}^{+\infty} (x - \mu)^2 \phi(x) \mathrm{d}x - \sigma^2 \right).$$

The maximum entropy probability density function meets Euler-Lagrange equation

$$\ln \phi(x) + 1 + \lambda_1 + \lambda_2 x + \lambda_3 (x - \mu)^2 = 0$$

and has the form $\phi(x) = \exp(-1 - \lambda_1 - \lambda_2 x - \lambda_3 (x - \mu)^2)$. Substituting it into the constraints, we get

$$\phi^*(x) = \frac{1}{\sigma\sqrt{2\pi}} \exp\left(-\frac{(x - \mu)^2}{2\sigma^2}\right), \quad x \in \Re$$

which is just the normally distributed random variable, and the maximum entropy is $H[\xi^*] = 1/2 + \ln\sqrt{2\pi}\sigma$.

A.10   Conditional Probability

We consider the probability of an event A after it has been learned that some other event B has occurred. This new probability of A is called the conditional probability of A given B.

Definition A.17 *Let $(\Omega, \mathcal{A}, \mathrm{Pr})$ be a probability space, and $A, B \in \mathcal{A}$. Then the conditional probability of A given B is defined by*

$$\mathrm{Pr}\{A|B\} = \frac{\mathrm{Pr}\{A \cap B\}}{\mathrm{Pr}\{B\}} \tag{A.35}$$

provided that $\mathrm{Pr}\{B\} > 0$.

Theorem A.21 *Let $(\Omega, \mathcal{A}, \mathrm{Pr})$ be a probability space, and B an event with $\mathrm{Pr}\{B\} > 0$. Then $\mathrm{Pr}\{\cdot|B\}$ defined by (A.35) is a probability measure, and $(\Omega, \mathcal{A}, \mathrm{Pr}\{\cdot|B\})$ is a probability space.*

Proof: It is sufficient to prove that $\mathrm{Pr}\{\cdot|B\}$ satisfies the normality, nonnegativity and countable additivity axioms. At first, we have

$$\mathrm{Pr}\{\Omega|B\} = \frac{\mathrm{Pr}\{\Omega \cap B\}}{\mathrm{Pr}\{B\}} = \frac{\mathrm{Pr}\{B\}}{\mathrm{Pr}\{B\}} = 1.$$

Secondly, for any $A \in \mathcal{A}$, the set function $\mathrm{Pr}\{A|B\}$ is nonnegative. Finally, for any countable sequence $\{A_i\}$ of mutually disjoint events, we have

$$\mathrm{Pr}\left\{\bigcup_{i=1}^{\infty} A_i | B\right\} = \frac{\mathrm{Pr}\left\{\left(\bigcup_{i=1}^{\infty} A_i\right) \cap B\right\}}{\mathrm{Pr}\{B\}} = \frac{\sum_{i=1}^{\infty} \mathrm{Pr}\{A_i \cap B\}}{\mathrm{Pr}\{B\}} = \sum_{i=1}^{\infty} \mathrm{Pr}\{A_i | B\}.$$

Thus $\mathrm{Pr}\{\cdot|B\}$ is a probability measure. Furthermore, $(\Omega, \mathcal{A}, \mathrm{Pr}\{\cdot|B\})$ is a probability space.

Theorem A.22 *(Bayes Formula) Let the events $A_1, A_2, \cdots, A_n$ form a partition of the space Ω such that $\Pr\{A_i\} > 0$ for $i = 1, 2, \cdots, n$, and let B be an event with $\Pr\{B\} > 0$. Then we have*

$$\Pr\{A_k|B\} = \frac{\Pr\{A_k\}\Pr\{B|A_k\}}{\sum\limits_{i=1}^{n}\Pr\{A_i\}\Pr\{B|A_i\}} \tag{A.36}$$

for $k = 1, 2, \cdots, n$.

Proof: Since $A_1, A_2, \cdots, A_n$ form a partition of the space Ω, we have

$$\Pr\{B\} = \sum_{i=1}^{n}\Pr\{A_i \cap B\} = \sum_{i=1}^{n}\Pr\{A_i\}\Pr\{B|A_i\}$$

which is also called the *formula for total probability*. Thus, for any k, we have

$$\Pr\{A_k|B\} = \frac{\Pr\{A_k \cap B\}}{\Pr\{B\}} = \frac{\Pr\{A_k\}\Pr\{B|A_k\}}{\sum\limits_{i=1}^{n}\Pr\{A_i\}\Pr\{B|A_i\}}.$$

The theorem is proved.

Remark A.1: Especially, let A and B be two events with $\Pr\{A\} > 0$ and $\Pr\{B\} > 0$. Then A and A^c form a partition of the space Ω, and the *Bayes formula* is

$$\Pr\{A|B\} = \frac{\Pr\{A\}\Pr\{B|A\}}{\Pr\{B\}}. \tag{A.37}$$

Remark A.2: In statistical applications, the events $A_1, A_2, \cdots, A_n$ are often called *hypotheses*. Furthermore, for each i, the $\Pr\{A_i\}$ is called the *prior* probability of A_i, and $\Pr\{A_i|B\}$ is called the *posterior* probability of A_i after the occurrence of event B.

Example A.18: Let ξ be an exponentially distributed random variable with expected value β. Then for any real numbers $a > 0$ and $x > 0$, the conditional probability of $\xi \geq a + x$ given $\xi \geq a$ is

$$\Pr\{\xi \geq a + x|\xi \geq a\} = \exp(-x/\beta) = \Pr\{\xi \geq x\}$$

which means that the conditional probability is identical to the original probability. This is the so-called memoryless property of exponential distribution. In other words, it is as good as new if it is functioning on inspection.

Definition A.18 *Let ξ be a random variable on $(\Omega, \mathcal{A}, \Pr)$. A conditional random variable of ξ given B is a measurable function $\xi|_B$ from the conditional probability space $(\Omega, \mathcal{A}, \Pr\{\cdot|B\})$ to the set of real numbers such that*

$$\xi|_B(\omega) \equiv \xi(\omega), \quad \forall \omega \in \Omega. \tag{A.38}$$

Definition A.19 *The conditional probability distribution* $\Phi\colon \Re \to [0,1]$ *of a random variable ξ given B is defined by*

$$\Phi(x|B) = \Pr\{\xi \le x|B\} \tag{A.39}$$

provided that $\Pr\{B\} > 0$.

Definition A.20 *The conditional probability density function ϕ of a random variable ξ given B is a nonnegative function such that*

$$\Phi(x|B) = \int_{-\infty}^{x} \phi(y|B)\mathrm{d}y, \quad \forall x \in \Re \tag{A.40}$$

where $\Phi(x|B)$ is the conditional probability distribution of ξ given B.

Example A.19: Let (ξ, η) be a random vector with joint probability density function ψ. Then the marginal probability density functions of ξ and η are

$$f(x) = \int_{-\infty}^{+\infty} \psi(x,y)\mathrm{d}y, \quad g(y) = \int_{-\infty}^{+\infty} \psi(x,y)\mathrm{d}x,$$

respectively. Furthermore, we have

$$\Pr\{\xi \le x, \eta \le y\} = \int_{-\infty}^{x}\int_{-\infty}^{y} \psi(r,t)\mathrm{d}r\mathrm{d}t = \int_{-\infty}^{y}\left[\int_{-\infty}^{x} \frac{\psi(r,t)}{g(t)}\mathrm{d}r\right]g(t)\mathrm{d}t$$

which implies that the conditional probability distribution of ξ given $\eta = y$ is

$$\Phi(x|\eta = y) = \int_{-\infty}^{x} \frac{\psi(r,y)}{g(y)}\mathrm{d}r, \quad \text{a.s.} \tag{A.41}$$

and the conditional probability density function of ξ given $\eta = y$ is

$$\phi(x|\eta = y) = \frac{\psi(x,y)}{g(y)} = \frac{\psi(x,y)}{\displaystyle\int_{-\infty}^{+\infty} \psi(x,y)\mathrm{d}x}, \quad \text{a.s.} \tag{A.42}$$

Note that (A.41) and (A.42) are defined only for $g(y) \ne 0$. In fact, the set $\{y|g(y) = 0\}$ has probability 0. Especially, if ξ and η are independent random variables, then $\psi(x,y) = f(x)g(y)$ and $\phi(x|\eta = y) = f(x)$.

Definition A.21 *Let ξ be a random variable. Then the conditional expected value of ξ given B is defined by*

$$E[\xi|B] = \int_{0}^{+\infty} \Pr\{\xi \ge r|B\}\mathrm{d}r - \int_{-\infty}^{0} \Pr\{\xi \le r|B\}\mathrm{d}r \tag{A.43}$$

provided that at least one of the two integrals is finite.

Appendix B

Credibility Theory

The concept of fuzzy set was initiated by Zadeh [229] via membership function in 1965. In order to measure a fuzzy event, Zadeh [232] proposed the concept of possibility measure. Although possibility measure has been widely used, it has no self-duality property. However, a self-dual measure is absolutely needed in both theory and practice. In order to define a self-dual measure, Liu and Liu [121] presented the concept of credibility measure. In addition, a sufficient and necessary condition for credibility measure was given by Li and Liu [94]. Credibility theory, founded by Liu [124] in 2004 and refined by Liu [127] in 2007, is a branch of mathematics for studying the behavior of fuzzy phenomena.

The emphasis in this appendix is mainly on credibility measure, credibility space, fuzzy variable, membership function, credibility distribution, independence, expected value, variance, moments, critical values, entropy and conditional credibility.

B.1 Credibility Space

Let Θ be a nonempty set, and $\mathcal{P}$ the power set of Θ (i.e., the larggest σ-algebra over Θ). Each element in $\mathcal{P}$ is called an *event*. In order to present an axiomatic definition of credibility, it is necessary to assign to each event A a number $\mathrm{Cr}\{A\}$ which indicates the credibility that A will occur. In order to ensure that the number $\mathrm{Cr}\{A\}$ has certain mathematical properties which we intuitively expect a credibility to have, we accept the following four axioms:

Axiom 1. *(Normality)* $\mathrm{Cr}\{\Theta\} = 1$.

Axiom 2. *(Monotonicity)* $\mathrm{Cr}\{A\} \leq \mathrm{Cr}\{B\}$ *whenever* $A \subset B$.

Axiom 3. *(Self-Duality)* $\mathrm{Cr}\{A\} + \mathrm{Cr}\{A^c\} = 1$ *for any event A.*

Axiom 4. *(Maximality)* $\mathrm{Cr}\{\cup_i A_i\} = \sup_i \mathrm{Cr}\{A_i\}$ *for any events $\{A_i\}$ with* $\sup_i \mathrm{Cr}\{A_i\} < 0.5$.

Definition B.1 *(Liu and Liu [121]) The set function* Cr *is called a credibility measure if it satisfies the normality, monotonicity, self-duality, and maximality axioms.*

Example B.1: Let $\Theta = \{\theta_1, \theta_2\}$. For this case, there are only four events: $\emptyset, \{\theta_1\}, \{\theta_2\}, \Theta$. Define $\mathrm{Cr}\{\emptyset\} = 0$, $\mathrm{Cr}\{\theta_1\} = 0.7$, $\mathrm{Cr}\{\theta_2\} = 0.3$, and $\mathrm{Cr}\{\Theta\} = 1$. Then the set function Cr is a credibility measure because it satisfies the four axioms.

Example B.2: Let Θ be a nonempty set. Define $\mathrm{Cr}\{\emptyset\} = 0$, $\mathrm{Cr}\{\Theta\} = 1$ and $\mathrm{Cr}\{A\} = 1/2$ for any subset A (excluding $\emptyset$ and Θ). Then the set function Cr is a credibility measure.

Example B.3: Let μ be a nonnegative function on Θ (for example, the set of real numbers) such that

$$\sup_{x \in \Theta} \mu(x) = 1. \tag{B.1}$$

Then the set function

$$\mathrm{Cr}\{A\} = \frac{1}{2} \left(\sup_{x \in A} \mu(x) + 1 - \sup_{x \in A^c} \mu(x) \right) \tag{B.2}$$

is a credibility measure on Θ.

Theorem B.1 *Let Θ be a nonempty set, $\mathcal{P}$ the power set of Θ, and* Cr *the credibility measure. Then $\mathrm{Cr}\{\emptyset\} = 0$ and $0 \leq \mathrm{Cr}\{A\} \leq 1$ for any $A \in \mathcal{P}$.*

Proof: It follows from Axioms 1 and 3 that $\mathrm{Cr}\{\emptyset\} = 1 - \mathrm{Cr}\{\Theta\} = 1 - 1 = 0$. Since $\emptyset \subset A \subset \Theta$, we have $0 \leq \mathrm{Cr}\{A\} \leq 1$ by using Axiom 2.

Theorem B.2 *Let Θ be a nonempty set, $\mathcal{P}$ the power set of Θ, and* Cr *the credibility measure. Then for any $A, B \in \mathcal{P}$, we have*

$$\mathrm{Cr}\{A \cup B\} = \mathrm{Cr}\{A\} \vee \mathrm{Cr}\{B\} \quad \text{if } \mathrm{Cr}\{A \cup B\} \leq 0.5, \tag{B.3}$$

$$\mathrm{Cr}\{A \cap B\} = \mathrm{Cr}\{A\} \wedge \mathrm{Cr}\{B\} \quad \text{if } \mathrm{Cr}\{A \cap B\} \geq 0.5. \tag{B.4}$$

The above equations hold for not only finite number of events but also infinite number of events.

Proof: If $\mathrm{Cr}\{A \cup B\} < 0.5$, then $\mathrm{Cr}\{A\} \vee \mathrm{Cr}\{B\} < 0.5$ by using Axiom 2. Thus the equation (B.3) follows immediately from Axiom 4. If $\mathrm{Cr}\{A \cup B\} = 0.5$ and (B.3) does not hold, then we have $\mathrm{Cr}\{A\} \vee \mathrm{Cr}\{B\} < 0.5$. It follows from Axiom 4 that

$$\mathrm{Cr}\{A \cup B\} = \mathrm{Cr}\{A\} \vee \mathrm{Cr}\{B\} < 0.5.$$

A contradiction proves (B.3). Next we prove (B.4). Since $\text{Cr}\{A \cap B\} \geq 0.5$, we have $\text{Cr}\{A^c \cup B^c\} \leq 0.5$ by the self-duality. Thus

$$\text{Cr}\{A \cap B\} = 1 - \text{Cr}\{A^c \cup B^c\} = 1 - \text{Cr}\{A^c\} \vee \text{Cr}\{B^c\}$$
$$= (1 - \text{Cr}\{A^c\}) \wedge (1 - \text{Cr}\{B^c\}) = \text{Cr}\{A\} \wedge \text{Cr}\{B\}.$$

The theorem is proved.

Theorem B.3 *Let* Θ *be a nonempty set,* $\mathcal{P}$ *the power set of* Θ*, and* Cr *the credibility measure. Then for any* $A, B \in \mathcal{P}$*, we have*

$$\text{Cr}\{A \cup B\} = \text{Cr}\{A\} \vee \text{Cr}\{B\} \quad \text{if } \text{Cr}\{A\} + \text{Cr}\{B\} < 1, \qquad (\text{B.5})$$

$$\text{Cr}\{A \cap B\} = \text{Cr}\{A\} \wedge \text{Cr}\{B\} \quad \text{if } \text{Cr}\{A\} + \text{Cr}\{B\} > 1. \qquad (\text{B.6})$$

Proof: Suppose $\text{Cr}\{A\} + \text{Cr}\{B\} < 1$. Then there exists at least one term less than 0.5, say $\text{Cr}\{B\} < 0.5$. If $\text{Cr}\{A\} < 0.5$ also holds, then the equation (B.3) follows immediately from Axiom 4. If $\text{Cr}\{A\} \geq 0.5$, then by using Theorem B.2, we obtain

$$\text{Cr}\{A\} = \text{Cr}\{A \cup (B \cap B^c)\} = \text{Cr}\{(A \cup B) \cap (A \cup B^c)\} = \text{Cr}\{A \cup B\} \wedge \text{Cr}\{A \cup B^c\}.$$

On the other hand, we have

$$\text{Cr}\{A\} < 1 - \text{Cr}\{B\} = \text{Cr}\{B^c\} \leq \text{Cr}\{A \cup B^c\}.$$

Hence we must have $\text{Cr}\{A \cup B\} = \text{Cr}\{A\} = \text{Cr}\{A\} \vee \text{Cr}\{B\}$. The equation (B.5) is proved. Next we suppose $\text{Cr}\{A\} + \text{Cr}\{B\} > 1$. Then $\text{Cr}\{A^c\} + \text{Cr}\{B^c\} < 1$. It follows from (B.5) that

$$\text{Cr}\{A \cap B\} = 1 - \text{Cr}\{A^c \cup B^c\} = 1 - \text{Cr}\{A^c\} \vee \text{Cr}\{B^c\}$$
$$= (1 - \text{Cr}\{A^c\}) \wedge (1 - \text{Cr}\{B^c\}) = \text{Cr}\{A\} \wedge \text{Cr}\{B\}.$$

The theorem is proved.

Credibility Subadditivity Theorem

Theorem B.4 *(Liu [124], Credibility Subadditivity Theorem) The credibility measure is subadditive. That is,*

$$\text{Cr}\{A \cup B\} \leq \text{Cr}\{A\} + \text{Cr}\{B\} \qquad (\text{B.7})$$

for any events A *and* B*. In fact, credibility measure is not only finitely subadditive but also countably subadditive.*

Proof: The argument breaks down into three cases. Case 1: $\mathrm{Cr}\{A\} < 0.5$ and $\mathrm{Cr}\{B\} < 0.5$. It follows from Axiom 4 that

$$\mathrm{Cr}\{A \cup B\} = \mathrm{Cr}\{A\} \vee \mathrm{Cr}\{B\} \leq \mathrm{Cr}\{A\} + \mathrm{Cr}\{B\}.$$

Case 2: $\mathrm{Cr}\{A\} \geq 0.5$. For this case, by using Axioms 2 and 3, we have $\mathrm{Cr}\{A^c\} \leq 0.5$ and $\mathrm{Cr}\{A \cup B\} \geq \mathrm{Cr}\{A\} \geq 0.5$. Then

$$\begin{aligned}
\mathrm{Cr}\{A^c\} &= \mathrm{Cr}\{A^c \cap B\} \vee \mathrm{Cr}\{A^c \cap B^c\} \\
&\leq \mathrm{Cr}\{A^c \cap B\} + \mathrm{Cr}\{A^c \cap B^c\} \\
&\leq \mathrm{Cr}\{B\} + \mathrm{Cr}\{A^c \cap B^c\}.
\end{aligned}$$

Applying this inequality, we obtain

$$\begin{aligned}
\mathrm{Cr}\{A\} + \mathrm{Cr}\{B\} &= 1 - \mathrm{Cr}\{A^c\} + \mathrm{Cr}\{B\} \\
&\geq 1 - \mathrm{Cr}\{B\} - \mathrm{Cr}\{A^c \cap B^c\} + \mathrm{Cr}\{B\} \\
&= 1 - \mathrm{Cr}\{A^c \cap B^c\} \\
&= \mathrm{Cr}\{A \cup B\}.
\end{aligned}$$

Case 3: $\mathrm{Cr}\{B\} \geq 0.5$. This case may be proved by a similar process of Case 2. The theorem is proved.

Remark B.1: For any events A and B, it follows from the credibility subadditivity theorem that the credibility measure is null-additive, i.e., $\mathrm{Cr}\{A \cup B\} = \mathrm{Cr}\{A\} + \mathrm{Cr}\{B\}$ if either $\mathrm{Cr}\{A\} = 0$ or $\mathrm{Cr}\{B\} = 0$.

Theorem B.5 *Let $\{B_i\}$ be a decreasing sequence of events with $\mathrm{Cr}\{B_i\} \to 0$ as $i \to \infty$. Then for any event A, we have*

$$\lim_{i \to \infty} \mathrm{Cr}\{A \cup B_i\} = \lim_{i \to \infty} \mathrm{Cr}\{A \backslash B_i\} = \mathrm{Cr}\{A\}. \tag{B.8}$$

Proof: It follows from the monotonicity axiom and credibility subadditivity theorem that

$$\mathrm{Cr}\{A\} \leq \mathrm{Cr}\{A \cup B_i\} \leq \mathrm{Cr}\{A\} + \mathrm{Cr}\{B_i\}$$

for each i. Thus we get $\mathrm{Cr}\{A \cup B_i\} \to \mathrm{Cr}\{A\}$ by using $\mathrm{Cr}\{B_i\} \to 0$. Since $(A \backslash B_i) \subset A \subset ((A \backslash B_i) \cup B_i)$, we have

$$\mathrm{Cr}\{A \backslash B_i\} \leq \mathrm{Cr}\{A\} \leq \mathrm{Cr}\{A \backslash B_i\} + \mathrm{Cr}\{B_i\}.$$

Hence $\mathrm{Cr}\{A \backslash B_i\} \to \mathrm{Cr}\{A\}$ by using $\mathrm{Cr}\{B_i\} \to 0$.

Credibility Semicontinuity Law

Generally speaking, the credibility measure is neither lower semicontinuous nor upper semicontinuous. However, we have the following credibility semicontinuity law.

Theorem B.6 *(Liu [124], Credibility Semicontinuity Law) For any events $A_1, A_2, \cdots$, we have*

$$\lim_{i \to \infty} \mathrm{Cr}\{A_i\} = \mathrm{Cr}\left\{\lim_{i \to \infty} A_i\right\} \tag{B.9}$$

if one of the following conditions is satisfied:
(a) $\mathrm{Cr}\{A\} \leq 0.5$ and $A_i \uparrow A$;   *(b) $\lim_{i \to \infty} \mathrm{Cr}\{A_i\} < 0.5$ and $A_i \uparrow A$;*
(c) $\mathrm{Cr}\{A\} \geq 0.5$ and $A_i \downarrow A$;   *(d) $\lim_{i \to \infty} \mathrm{Cr}\{A_i\} > 0.5$ and $A_i \downarrow A$.*

Proof: (a) Since $\mathrm{Cr}\{A\} \leq 0.5$, we have $\mathrm{Cr}\{A_i\} \leq 0.5$ for each i. It follows from Axiom 4 that

$$\mathrm{Cr}\{A\} = \mathrm{Cr}\{\cup_i A_i\} = \sup_i \mathrm{Cr}\{A_i\} = \lim_{i \to \infty} \mathrm{Cr}\{A_i\}.$$

(b) Since $\lim_{i \to \infty} \mathrm{Cr}\{A_i\} < 0.5$, we have $\sup_i \mathrm{Cr}\{A_i\} < 0.5$. It follows from Axiom 4 that

$$\mathrm{Cr}\{A\} = \mathrm{Cr}\{\cup_i A_i\} = \sup_i \mathrm{Cr}\{A_i\} = \lim_{i \to \infty} \mathrm{Cr}\{A_i\}.$$

(c) Since $\mathrm{Cr}\{A\} \geq 0.5$ and $A_i \downarrow A$, it follows from the self-duality of credibility measure that $\mathrm{Cr}\{A^c\} \leq 0.5$ and $A_i^c \uparrow A^c$. Thus

$$\lim_{i \to \infty} \mathrm{Cr}\{A_i\} = 1 - \lim_{i \to \infty} \mathrm{Cr}\{A_i^c\} = 1 - \mathrm{Cr}\{A^c\} = \mathrm{Cr}\{A\}.$$

(d) Since $\lim_{i \to \infty} \mathrm{Cr}\{A_i\} > 0.5$ and $A_i \downarrow A$, it follows from the self-duality of credibility measure that

$$\lim_{i \to \infty} \mathrm{Cr}\{A_i^c\} = \lim_{i \to \infty} \left(1 - \mathrm{Cr}\{A_i\}\right) < 0.5$$

and $A_i^c \uparrow A^c$. Thus $\mathrm{Cr}\{A_i\} = 1 - \mathrm{Cr}\{A_i^c\} \to 1 - \mathrm{Cr}\{A^c\} = \mathrm{Cr}\{A\}$ as $i \to \infty$. The theorem is proved.

Credibility Asymptotic Theorem

Theorem B.7 *(Credibility Asymptotic Theorem) For any events $A_1, A_2, \cdots$, we have*

$$\lim_{i \to \infty} \mathrm{Cr}\{A_i\} \geq 0.5, \quad if \ A_i \uparrow \Theta, \tag{B.10}$$

$$\lim_{i \to \infty} \mathrm{Cr}\{A_i\} \leq 0.5, \quad if \ A_i \downarrow \emptyset. \tag{B.11}$$

Proof: Assume $A_i \uparrow \Theta$. If $\lim_{i \to \infty} \mathrm{Cr}\{A_i\} < 0.5$, it follows from the credibility semicontinuity law that

$$\mathrm{Cr}\{\Theta\} = \lim_{i \to \infty} \mathrm{Cr}\{A_i\} < 0.5$$

which is in contradiction with $\mathrm{Cr}\{\Theta\} = 1$. The first inequality is proved. The second one may be verified similarly.

Credibility Extension Theorem

Suppose that the credibility of each singleton is given. Is the credibility measure fully and uniquely determined? This subsection will answer the question.

Theorem B.8 *Suppose that Θ is a nonempty set. If Cr is a credibility measure, then we have*

$$\sup_{\theta \in \Theta} \mathrm{Cr}\{\theta\} \geq 0.5,$$

$$\mathrm{Cr}\{\theta^*\} + \sup_{\theta \neq \theta^*} \mathrm{Cr}\{\theta\} = 1 \ \text{if} \ \mathrm{Cr}\{\theta^*\} \geq 0.5. \tag{B.12}$$

We will call (B.12) the credibility extension condition.

Proof: If $\sup \mathrm{Cr}\{\theta\} < 0.5$, then by using Axiom 4, we have

$$1 = \mathrm{Cr}\{\Theta\} = \sup_{\theta \in \Theta} \mathrm{Cr}\{\theta\} < 0.5.$$

This contradiction proves $\sup \mathrm{Cr}\{\theta\} \geq 0.5$. We suppose that $\theta^* \in \Theta$ is a point with $\mathrm{Cr}\{\theta^*\} \geq 0.5$. It follows from Axioms 3 and 4 that $\mathrm{Cr}\{\Theta \setminus \{\theta^*\}\} \leq 0.5$, and

$$\mathrm{Cr}\{\Theta \setminus \{\theta^*\}\} = \sup_{\theta \neq \theta^*} \mathrm{Cr}\{\theta\}.$$

Hence the second formula of (B.12) is true by the self-duality of credibility measure.

Theorem B.9 *(Li and Liu [94], Credibility Extension Theorem) Suppose that Θ is a nonempty set, and $\mathrm{Cr}\{\theta\}$ is a nonnegative function on Θ satisfying the credibility extension condition (B.12). Then $\mathrm{Cr}\{\theta\}$ has a unique extension to a credibility measure as follows,*

$$\mathrm{Cr}\{A\} = \begin{cases} \sup_{\theta \in A} \mathrm{Cr}\{\theta\}, & \text{if} \ \sup_{\theta \in A} \mathrm{Cr}\{\theta\} < 0.5 \\ 1 - \sup_{\theta \in A^c} \mathrm{Cr}\{\theta\}, & \text{if} \ \sup_{\theta \in A} \mathrm{Cr}\{\theta\} \geq 0.5. \end{cases} \tag{B.13}$$

Proof: We first prove that the set function $\mathrm{Cr}\{A\}$ defined by (B.13) is a credibility measure.

STEP 1: By the credibility extension condition $\sup_{\theta \in \Theta} \mathrm{Cr}\{\theta\} \geq 0.5$, we have

$$\mathrm{Cr}\{\Theta\} = 1 - \sup_{\theta \in \emptyset} \mathrm{Cr}\{\theta\} = 1 - 0 = 1.$$

STEP 2: If $A \subset B$, then $B^c \subset A^c$. The proof breaks down into two cases. Case 1: $\sup_{\theta \in A} \mathrm{Cr}\{\theta\} < 0.5$. For this case, we have

$$\mathrm{Cr}\{A\} = \sup_{\theta \in A} \mathrm{Cr}\{\theta\} \leq \sup_{\theta \in B} \mathrm{Cr}\{\theta\} \leq \mathrm{Cr}\{B\}.$$

Case 2: $\sup_{\theta \in A} \mathrm{Cr}\{\theta\} \geq 0.5$. For this case, we have $\sup_{\theta \in B} \mathrm{Cr}\{\theta\} \geq 0.5$, and

$$\mathrm{Cr}\{A\} = 1 - \sup_{\theta \in A^c} \mathrm{Cr}\{\theta\} \leq 1 - \sup_{\theta \in B^c} \mathrm{Cr}\{\theta\} = \mathrm{Cr}\{B\}.$$

STEP 3: In order to prove $\mathrm{Cr}\{A\} + \mathrm{Cr}\{A^c\} = 1$, the argument breaks down into two cases. Case 1: $\sup_{\theta \in A} \mathrm{Cr}\{\theta\} < 0.5$. For this case, we have $\sup_{\theta \in A^c} \mathrm{Cr}\{\theta\} \geq 0.5$. Thus,

$$\mathrm{Cr}\{A\} + \mathrm{Cr}\{A^c\} = \sup_{\theta \in A} \mathrm{Cr}\{\theta\} + 1 - \sup_{\theta \in A} \mathrm{Cr}\{\theta\} = 1.$$

Case 2: $\sup_{\theta \in A} \mathrm{Cr}\{\theta\} \geq 0.5$. For this case, we have $\sup_{\theta \in A^c} \mathrm{Cr}\{\theta\} \leq 0.5$, and

$$\mathrm{Cr}\{A\} + \mathrm{Cr}\{A^c\} = 1 - \sup_{\theta \in A^c} \mathrm{Cr}\{\theta\} + \sup_{\theta \in A^c} \mathrm{Cr}\{\theta\} = 1.$$

STEP 4: For any collection $\{A_i\}$ with $\sup_i \mathrm{Cr}\{A_i\} < 0.5$, we have

$$\mathrm{Cr}\{\cup_i A_i\} = \sup_{\theta \in \cup_i A_i} \mathrm{Cr}\{\theta\} = \sup_i \sup_{\theta \in A_i} \mathrm{Cr}\{\theta\} = \sup_i \mathrm{Cr}\{A_i\}.$$

Thus Cr is a credibility measure because it satisfies the four axioms.

Finally, let us prove the uniqueness. Assume that Cr_1 and Cr_2 are two credibility measures such that $\mathrm{Cr}_1\{\theta\} = \mathrm{Cr}_2\{\theta\}$ for each $\theta \in \Theta$. Let us prove that $\mathrm{Cr}_1\{A\} = \mathrm{Cr}_2\{A\}$ for any event A. The argument breaks down into three cases. Case 1: $\mathrm{Cr}_1\{A\} < 0.5$. For this case, it follows from Axiom 4 that

$$\mathrm{Cr}_1\{A\} = \sup_{\theta \in A} \mathrm{Cr}_1\{\theta\} = \sup_{\theta \in A} \mathrm{Cr}_2\{\theta\} = \mathrm{Cr}_2\{A\}.$$

Case 2: $\mathrm{Cr}_1\{A\} > 0.5$. For this case, we have $\mathrm{Cr}_1\{A^c\} < 0.5$. It follows from the first case that $\mathrm{Cr}_1\{A^c\} = \mathrm{Cr}_2\{A^c\}$ which implies $\mathrm{Cr}_1\{A\} = \mathrm{Cr}_2\{A\}$. Case 3: $\mathrm{Cr}_1\{A\} = 0.5$. For this case, we have $\mathrm{Cr}_1\{A^c\} = 0.5$, and

$$\mathrm{Cr}_2\{A\} \geq \sup_{\theta \in A} \mathrm{Cr}_2\{\theta\} = \sup_{\theta \in A} \mathrm{Cr}_1\{\theta\} = \mathrm{Cr}_1\{A\} = 0.5,$$

$$\mathrm{Cr}_2\{A^c\} \geq \sup_{\theta \in A^c} \mathrm{Cr}_2\{\theta\} = \sup_{\theta \in A^c} \mathrm{Cr}_1\{\theta\} = \mathrm{Cr}_1\{A^c\} = 0.5.$$

Hence $\mathrm{Cr}_2\{A\} = 0.5 = \mathrm{Cr}_1\{A\}$. The uniqueness is proved.

Credibility Space

Definition B.2 *Let Θ be a nonempty set, $\mathcal{P}$ the power set of Θ, and Cr a credibility measure. Then the triplet $(\Theta, \mathcal{P}, \text{Cr})$ is called a credibility space.*

Example B.4: The triplet $(\Theta, \mathcal{P}, \text{Cr})$ is a credibility space if

$$\Theta = \{\theta_1, \theta_2, \cdots\},\ \text{Cr}\{\theta_i\} \equiv 1/2 \ \text{for } i = 1, 2, \cdots \qquad (\text{B.14})$$

Note that the credibility measure is produced by the credibility extension theorem as follows,

$$\text{Cr}\{A\} = \begin{cases} 0, & \text{if } A = \emptyset \\ 1, & \text{if } A = \Theta \\ 1/2, & \text{otherwise.} \end{cases}$$

Example B.5: The triplet $(\Theta, \mathcal{P}, \text{Cr})$ is a credibility space if

$$\Theta = \{\theta_1, \theta_2, \cdots\},\ \text{Cr}\{\theta_i\} = i/(2i+1) \ \text{for } i = 1, 2, \cdots \qquad (\text{B.15})$$

By using the credibility extension theorem, we obtain the following credibility measure,

$$\text{Cr}\{A\} = \begin{cases} \displaystyle\sup_{\theta_i \in A} \frac{i}{2i+1}, & \text{if } A \text{ is finite} \\ 1 - \displaystyle\sup_{\theta_i \in A^c} \frac{i}{2i+1}, & \text{if } A \text{ is infinite.} \end{cases}$$

Example B.6: The triplet $(\Theta, \mathcal{P}, \text{Cr})$ is a credibility space if

$$\Theta = \{\theta_1, \theta_2, \cdots\},\ \text{Cr}\{\theta_1\} = 1/2,\ \text{Cr}\{\theta_i\} = 1/i \ \text{for } i = 2, 3, \cdots \qquad (\text{B.16})$$

For this case, the credibility measure is

$$\text{Cr}\{A\} = \begin{cases} \displaystyle\sup_{\theta_i \in A} 1/i, & \text{if } A \text{ contains neither } \theta_1 \text{ nor } \theta_2 \\ 1/2, & \text{if } A \text{ contains only one of } \theta_1 \text{ and } \theta_2 \\ 1 - \displaystyle\sup_{\theta_i \in A^c} 1/i, & \text{if } A \text{ contains both } \theta_1 \text{ and } \theta_2. \end{cases}$$

Example B.7: The triplet $(\Theta, \mathcal{P}, \text{Cr})$ is a credibility space if

$$\Theta = [0, 1], \quad \text{Cr}\{\theta\} = \theta/2 \ \text{for } \theta \in \Theta. \qquad (\text{B.17})$$

For this case, the credibility measure is

$$\text{Cr}\{A\} = \begin{cases} \dfrac{1}{2} \displaystyle\sup_{\theta \in A} \theta, & \text{if } \displaystyle\sup_{\theta \in A} \theta < 1 \\ 1 - \dfrac{1}{2} \displaystyle\sup_{\theta \in A^c} \theta, & \text{if } \displaystyle\sup_{\theta \in A} \theta = 1. \end{cases}$$

Product Credibility Measure

Product credibility measure may be defined in multiple ways. This book accepts the following axiom.

Axiom 5. *(Product Credibility Axiom) Let Θ_k be nonempty sets on which Cr_k are credibility measures, $k = 1, 2, \cdots, n$, respectively, and $\Theta = \Theta_1 \times \Theta_2 \times \cdots \times \Theta_n$. Then*

$$\mathrm{Cr}\{(\theta_1, \theta_2, \cdots, \theta_n)\} = \mathrm{Cr}_1\{\theta_1\} \wedge \mathrm{Cr}_2\{\theta_2\} \wedge \cdots \wedge \mathrm{Cr}_n\{\theta_n\} \tag{B.18}$$

for each $(\theta_1, \theta_2, \cdots, \theta_n) \in \Theta$.

Theorem B.10 *(Product Credibility Theorem) Let Θ_k be nonempty sets on which Cr_k are the credibility measures, $k = 1, 2, \cdots, n$, respectively, and $\Theta = \Theta_1 \times \Theta_2 \times \cdots \times \Theta_n$. Then $\mathrm{Cr} = \mathrm{Cr}_1 \wedge \mathrm{Cr}_2 \wedge \cdots \wedge \mathrm{Cr}_n$ defined by Axiom 5 has a unique extension to a credibility measure on Θ as follows,*

$$\mathrm{Cr}\{A\} = \begin{cases} \displaystyle\sup_{(\theta_1, \theta_2 \cdots, \theta_n) \in A} \min_{1 \leq k \leq n} \mathrm{Cr}_k\{\theta_k\}, \\ \qquad if \ \displaystyle\sup_{(\theta_1, \theta_2, \cdots, \theta_n) \in A} \min_{1 \leq k \leq n} \mathrm{Cr}_k\{\theta_k\} < 0.5 \\ 1 - \displaystyle\sup_{(\theta_1, \theta_2, \cdots, \theta_n) \in A^c} \min_{1 \leq k \leq n} \mathrm{Cr}_k\{\theta_k\}, \\ \qquad if \ \displaystyle\sup_{(\theta_1, \theta_2, \cdots, \theta_n) \in A} \min_{1 \leq k \leq n} \mathrm{Cr}_k\{\theta_k\} \geq 0.5. \end{cases} \tag{B.19}$$

Proof: For each $\boldsymbol{\theta} = (\theta_1, \theta_2, \cdots, \theta_n) \in \Theta$, we have $\mathrm{Cr}\{\boldsymbol{\theta}\} = \mathrm{Cr}_1\{\theta_1\} \wedge \mathrm{Cr}_2\{\theta_2\} \wedge \cdots \wedge \mathrm{Cr}_n\{\theta_n\}$. Let us prove that $\mathrm{Cr}\{\boldsymbol{\theta}\}$ satisfies the credibility extension condition. Since $\sup_{\theta_k \in \Theta_k} \mathrm{Cr}\{\theta_k\} \geq 0.5$ for each k, we have

$$\sup_{\boldsymbol{\theta} \in \Theta} \mathrm{Cr}\{\boldsymbol{\theta}\} = \sup_{(\theta_1, \theta_2, \cdots, \theta_n) \in \Theta} \min_{1 \leq k \leq n} \mathrm{Cr}_k\{\theta_k\} \geq 0.5.$$

Now we suppose that $\boldsymbol{\theta}^* = (\theta_1^*, \theta_2^*, \cdots, \theta_n^*)$ is a point with $\mathrm{Cr}\{\boldsymbol{\theta}^*\} \geq 0.5$. Without loss of generality, let i be the index such that

$$\mathrm{Cr}\{\boldsymbol{\theta}^*\} = \min_{1 \leq k \leq n} \mathrm{Cr}_k\{\theta_k^*\} = \mathrm{Cr}_i\{\theta_i^*\}. \tag{B.20}$$

We also immediately have

$$\mathrm{Cr}_k\{\theta_k^*\} \geq 0.5, \quad k = 1, 2, \cdots, n; \tag{B.21}$$

$$\mathrm{Cr}_k\{\theta_k^*\} + \sup_{\theta_k \neq \theta_k^*} \mathrm{Cr}_k\{\theta_k\} = 1, \quad k = 1, 2, \cdots, n; \tag{B.22}$$

$$\sup_{\theta_i \neq \theta_i^*} \mathrm{Cr}_i\{\theta_i\} \geq \sup_{\theta_k \neq \theta_k^*} \mathrm{Cr}_k\{\theta_k\}, \quad k = 1, 2, \cdots, n; \tag{B.23}$$

$$\sup_{\theta_k \neq \theta_k^*} \mathrm{Cr}_k\{\theta_k\} \leq 0.5, \quad k = 1, \cdots, n. \tag{B.24}$$

It follows from (B.21) and (B.24) that

$$\begin{aligned}
\sup_{\boldsymbol{\theta} \neq \boldsymbol{\theta}^*} \mathrm{Cr}\{\boldsymbol{\theta}\} &= \sup_{(\theta_1, \theta_2, \cdots, \theta_n) \neq (\theta_1^*, \theta_2^*, \cdots, \theta_n^*)} \min_{1 \leq k \leq n} \mathrm{Cr}_k\{\theta_k\} \\
&\geq \sup_{\theta_i \neq \theta_i^*} \min_{1 \leq k \leq i-1} \mathrm{Cr}_k\{\theta_k^*\} \wedge \mathrm{Cr}_i\{\theta_i\} \wedge \min_{i+1 \leq k \leq n} \mathrm{Cr}_k\{\theta_k^*\} \\
&= \sup_{\theta_i \neq \theta_i^*} \mathrm{Cr}_i\{\theta_i\}.
\end{aligned}$$

We next suppose that

$$\sup_{\boldsymbol{\theta} \neq \boldsymbol{\theta}^*} \mathrm{Cr}\{\boldsymbol{\theta}\} > \sup_{\theta_i \neq \theta_i^*} \mathrm{Cr}_i\{\theta_i\}.$$

Then there is a point $(\theta_1', \theta_2', \cdots, \theta_n') \neq (\theta_1^*, \theta_2^*, \cdots, \theta_n^*)$ such that

$$\min_{1 \leq k \leq n} \mathrm{Cr}_k\{\theta_k'\} > \sup_{\theta_i \neq \theta_i^*} \mathrm{Cr}_i\{\theta_i\}.$$

Let j be one of the index such that $\theta_j' \neq \theta_j^*$. Then

$$\mathrm{Cr}_j\{\theta_j'\} > \sup_{\theta_i \neq \theta_i^*} \mathrm{Cr}_i\{\theta_i\}.$$

That is,

$$\sup_{\theta_j \neq \theta_j^*} \mathrm{Cr}_j\{\theta_j\} > \sup_{\theta_i \neq \theta_i^*} \mathrm{Cr}_i\{\theta_i\}$$

which is in contradiction with (B.23). Thus

$$\sup_{\boldsymbol{\theta} \neq \boldsymbol{\theta}^*} \mathrm{Cr}\{\boldsymbol{\theta}\} = \sup_{\theta_i \neq \theta_i^*} \mathrm{Cr}_i\{\theta_i\}. \tag{B.25}$$

It follows from (B.20), (B.22) and (B.25) that

$$\mathrm{Cr}\{\boldsymbol{\theta}^*\} + \sup_{\boldsymbol{\theta} \neq \boldsymbol{\theta}^*} \mathrm{Cr}\{\boldsymbol{\theta}\} = \mathrm{Cr}_i\{\theta_i^*\} + \sup_{\theta_i \neq \theta_i^*} \mathrm{Cr}_i\{\theta_i\} = 1.$$

Thus Cr satisfies the credibility extension condition. It follows from the credibility extension theorem that $\mathrm{Cr}\{A\}$ is just the unique extension of $\mathrm{Cr}\{\boldsymbol{\theta}\}$. The theorem is proved.

Definition B.3 *Let* $(\Theta_k, \mathcal{P}_k, \mathrm{Cr}_k), k = 1, 2, \cdots, n$ *be credibility spaces,* $\Theta = \Theta_1 \times \Theta_2 \times \cdots \times \Theta_n$ *and* $\mathrm{Cr} = \mathrm{Cr}_1 \wedge \mathrm{Cr}_2 \wedge \cdots \wedge \mathrm{Cr}_n$. *Then* $(\Theta, \mathcal{P}, \mathrm{Cr})$ *is called the product credibility space of* $(\Theta_k, \mathcal{P}_k, \mathrm{Cr}_k), k = 1, 2, \cdots, n$.

Theorem B.11 *Let* $(\Theta, \mathcal{P}, \mathrm{Cr})$ *be the product credibility space of* $(\Theta_k, \mathcal{P}_k, \mathrm{Cr}_k)$, $k = 1, 2, \cdots, n$. *Then for any* $A_k \in \mathcal{P}_k$, $k = 1, 2, \cdots, n$, *we have*

$$\mathrm{Cr}\{A_1 \times A_2 \times \cdots \times A_k\} = \mathrm{Cr}_1\{A_1\} \wedge \mathrm{Cr}_2\{A_2\} \wedge \cdots \wedge \mathrm{Cr}_n\{A_n\}.$$

Proof: We only prove the case of $n = 2$. If $\mathrm{Cr}_1\{A_1\} < 0.5$ or $\mathrm{Cr}_2\{A_2\} < 0.5$, then we have

$$\sup_{\theta_1 \in A_1} \mathrm{Cr}_1\{\theta_1\} < 0.5 \quad \text{or} \quad \sup_{\theta_2 \in A_2} \mathrm{Cr}_2\{\theta_2\} < 0.5.$$

It follows from

$$\sup_{(\theta_1,\theta_2) \in A_1 \times A_2} \mathrm{Cr}_1\{\theta_1\} \wedge \mathrm{Cr}_2\{\theta_2\} = \sup_{\theta_1 \in A_1} \mathrm{Cr}_1\{\theta_1\} \wedge \sup_{\theta_2 \in A_2} \mathrm{Cr}_2\{\theta_2\} < 0.5$$

that

$$\mathrm{Cr}\{A_1 \times A_2\} = \sup_{\theta_1 \in A_1} \mathrm{Cr}_1\{\theta_1\} \wedge \sup_{\theta_2 \in A_2} \mathrm{Cr}_2\{\theta_2\} = \mathrm{Cr}_1\{A_1\} \wedge \mathrm{Cr}_2\{A_2\}.$$

If $\mathrm{Cr}_1\{A_1\} \geq 0.5$ and $\mathrm{Cr}_2\{A_2\} \geq 0.5$, then we have

$$\sup_{\theta_1 \in A_1} \mathrm{Cr}_1\{\theta_1\} \geq 0.5 \quad \text{and} \quad \sup_{\theta_2 \in A_2} \mathrm{Cr}_2\{\theta_2\} \geq 0.5.$$

It follows from

$$\sup_{(\theta_1,\theta_2) \in A_1 \times A_2} \mathrm{Cr}_1\{\theta_1\} \wedge \mathrm{Cr}_2\{\theta_2\} = \sup_{\theta_1 \in A_1} \mathrm{Cr}_1\{\theta_1\} \wedge \sup_{\theta_2 \in A_2} \mathrm{Cr}_2\{\theta_2\} \geq 0.5$$

that

$$\begin{aligned}
\mathrm{Cr}\{A_1 \times A_2\} &= 1 - \sup_{(\theta_1,\theta_2) \notin A_1 \times A_2} \mathrm{Cr}_1\{\theta_1\} \wedge \mathrm{Cr}_2\{\theta_2\} \\
&= \left(1 - \sup_{\theta_1 \in A_1^c} \mathrm{Cr}_1\{\theta_1\}\right) \wedge \left(1 - \sup_{\theta_2 \in A_2^c} \mathrm{Cr}_2\{\theta_2\}\right) \\
&= \mathrm{Cr}_1\{A_1\} \wedge \mathrm{Cr}_2\{A_2\}.
\end{aligned}$$

The theorem is proved.

B.2 Fuzzy Variable

Definition B.4 *A fuzzy variable is a (measurable) function from a credibility space $(\Theta, \mathcal{P}, \mathrm{Cr})$ to the set of real numbers.*

Example B.8: Take $(\Theta, \mathcal{P}, \mathrm{Cr})$ to be $\{\theta_1, \theta_2\}$ with $\mathrm{Cr}\{\theta_1\} = \mathrm{Cr}\{\theta_2\} = 0.5$. Then the function

$$\xi(\theta) = \begin{cases} 0, & \text{if } \theta = \theta_1 \\ 1, & \text{if } \theta = \theta_2 \end{cases}$$

is a fuzzy variable.

Example B.9: Take $(\Theta, \mathcal{P}, \mathrm{Cr})$ to be the interval $[0, 1]$ with $\mathrm{Cr}\{\theta\} = \theta/2$ for each $\theta \in [0, 1]$. Then the identity function $\xi(\theta) = \theta$ is a fuzzy variable.

Example B.10: A crisp number c may be regarded as a special fuzzy variable. In fact, it is the constant function $\xi(\theta) \equiv c$ on the credibility space $(\Theta, \mathcal{P}, \mathrm{Cr})$.

Remark B.2: Since a fuzzy variable ξ is a function on a credibility space, for any set B of real numbers, the set

$$\{\xi \in B\} = \big\{\theta \in \Theta \mid \xi(\theta) \in B\big\} \tag{B.26}$$

is always an element in $\mathcal{P}$. In other words, the fuzzy variable ξ is always a measurable function and $\{\xi \in B\}$ is always an event.

Definition B.5 *Let ξ_1 and ξ_2 be fuzzy variables defined on the credibility space $(\Theta, \mathcal{P}, \mathrm{Cr})$. We say $\xi_1 = \xi_2$ if $\xi_1(\theta) = \xi_2(\theta)$ for almost all $\theta \in \Theta$.*

Fuzzy Vector

Definition B.6 *An n-dimensional fuzzy vector is defined as a function from a credibility space $(\Theta, \mathcal{P}, \mathrm{Cr})$ to the set of n-dimensional real vectors.*

Theorem B.12 *The vector $(\xi_1, \xi_2, \cdots, \xi_n)$ is a fuzzy vector if and only if $\xi_1, \xi_2, \cdots, \xi_n$ are fuzzy variables.*

Proof: Write $\boldsymbol{\xi} = (\xi_1, \xi_2, \cdots, \xi_n)$. Suppose that $\boldsymbol{\xi}$ is a fuzzy vector. Then $\xi_1, \xi_2, \cdots, \xi_n$ are functions from Θ to $\Re$. Thus $\xi_1, \xi_2, \cdots, \xi_n$ are fuzzy variables. Conversely, suppose that ξ_i are fuzzy variables defined on the credibility spaces $(\Theta_i, \mathcal{P}_i, \mathrm{Cr}_i)$, $i = 1, 2, \cdots, n$, respectively. It is clear that $(\xi_1, \xi_2, \cdots, \xi_n)$ is a function from the product credibility space $(\Theta, \mathcal{P}, \mathrm{Cr})$ to $\Re^n$, i.e.,

$$\boldsymbol{\xi}(\theta_1, \theta_2, \cdots, \theta_n) = (\xi_1(\theta_1), \xi_2(\theta_2), \cdots, \xi_n(\theta_n))$$

for all $(\theta_1, \theta_2, \cdots, \theta_n) \in \Theta$. Hence $\boldsymbol{\xi} = (\xi_1, \xi_2, \cdots, \xi_n)$ is a fuzzy vector.

Fuzzy Arithmetic

In this subsection, we will suppose that all fuzzy variables are defined on a common credibility space. Otherwise, we may embed them into the product credibility space.

Definition B.7 *Let $f : \Re^n \to \Re$ be a function, and $\xi_1, \xi_2, \cdots, \xi_n$ fuzzy variables on the credibility space $(\Theta, \mathcal{P}, \mathrm{Cr})$. Then $\xi = f(\xi_1, \xi_2, \cdots, \xi_n)$ is a fuzzy variable defined as*

$$\xi(\theta) = f(\xi_1(\theta), \xi_2(\theta), \cdots, \xi_n(\theta)) \tag{B.27}$$

for any $\theta \in \Theta$.

The reader may wonder whether $\xi(\theta_1, \theta_2, \cdots, \theta_n)$ defined by (B.27) is a fuzzy variable. The following theorem answers this question.

Theorem B.13 *Let $\boldsymbol{\xi}$ be an n-dimensional fuzzy vector, and $f : \Re^n \to \Re$ a function. Then $f(\boldsymbol{\xi})$ is a fuzzy variable.*

Proof: Since $f(\boldsymbol{\xi})$ is a function from a credibility space to the set of real numbers, it is a fuzzy variable.

B.3   Membership Function

Definition B.8 *Let ξ be a fuzzy variable defined on the credibility space $(\Theta, \mathcal{P}, \mathrm{Cr})$. Then its membership function is derived from the credibility measure by*

$$\mu(x) = (2\mathrm{Cr}\{\xi = x\}) \wedge 1, \quad x \in \Re. \tag{B.28}$$

Membership function represents the degree that the fuzzy variable ξ takes some prescribed value. How do we determine membership functions? There are several methods reported in the past literature. Anyway, the membership degree $\mu(x) = 0$ if x is an impossible point, and $\mu(x) = 1$ if x is the most possible point that ξ takes.

Example B.11: It is clear that a fuzzy variable has a unique membership function. However, a membership function may produce multiple fuzzy variables. For example, let $\Theta = \{\theta_1, \theta_2\}$ and $\mathrm{Cr}\{\theta_1\} = \mathrm{Cr}\{\theta_2\} = 0.5$. Then $(\Theta, \mathcal{P}, \mathrm{Cr})$ is a credibility space. We define

$$\xi_1(\theta) = \begin{cases} 0, & \text{if } \theta = \theta_1 \\ 1, & \text{if } \theta = \theta_2, \end{cases} \qquad \xi_2(\theta) = \begin{cases} 1, & \text{if } \theta = \theta_1 \\ 0, & \text{if } \theta = \theta_2. \end{cases}$$

It is clear that both of them are fuzzy variables and have the same membership function, $\mu(x) \equiv 1$ on $x = 0$ or 1.

Theorem B.14 *(Credibility Inversion Theorem) Let ξ be a fuzzy variable with membership function μ. Then for any set B of real numbers, we have*

$$\mathrm{Cr}\{\xi \in B\} = \frac{1}{2} \left(\sup_{x \in B} \mu(x) + 1 - \sup_{x \in B^c} \mu(x) \right). \tag{B.29}$$

Proof: If $\mathrm{Cr}\{\xi \in B\} \leq 0.5$, then by Axiom 2, we have $\mathrm{Cr}\{\xi = x\} \leq 0.5$ for each $x \in B$. It follows from Axiom 4 that

$$\mathrm{Cr}\{\xi \in B\} = \frac{1}{2} \left(\sup_{x \in B} (2\mathrm{Cr}\{\xi = x\} \wedge 1) \right) = \frac{1}{2} \sup_{x \in B} \mu(x). \tag{B.30}$$

The self-duality of credibility measure implies that $\mathrm{Cr}\{\xi \in B^c\} \geq 0.5$ and $\sup_{x \in B^c} \mathrm{Cr}\{\xi = x\} \geq 0.5$, i.e.,

$$\sup_{x \in B^c} \mu(x) = \sup_{x \in B^c} (2\mathrm{Cr}\{\xi = x\} \wedge 1) = 1. \tag{B.31}$$

It follows from (B.30) and (B.31) that (B.29) holds.

If $\mathrm{Cr}\{\xi \in B\} \geq 0.5$, then $\mathrm{Cr}\{\xi \in B^c\} \leq 0.5$. It follows from the first case that

$$\mathrm{Cr}\{\xi \in B\} = 1 - \mathrm{Cr}\{\xi \in B^c\} = 1 - \frac{1}{2}\left(\sup_{x \in B^c} \mu(x) + 1 - \sup_{x \in B} \mu(x)\right)$$

$$= \frac{1}{2}\left(\sup_{x \in B} \mu(x) + 1 - \sup_{x \in B^c} \mu(x)\right).$$

The theorem is proved.

Example B.12: Let ξ be a fuzzy variable with membership function μ. Then the following equations follow immediately from Theorem B.14:

$$\mathrm{Cr}\{\xi = x\} = \frac{1}{2}\left(\mu(x) + 1 - \sup_{y \neq x} \mu(y)\right), \quad \forall x \in \Re; \tag{B.32}$$

$$\mathrm{Cr}\{\xi \leq x\} = \frac{1}{2}\left(\sup_{y \leq x} \mu(y) + 1 - \sup_{y > x} \mu(y)\right), \quad \forall x \in \Re; \tag{B.33}$$

$$\mathrm{Cr}\{\xi \geq x\} = \frac{1}{2}\left(\sup_{y \geq x} \mu(y) + 1 - \sup_{y < x} \mu(y)\right), \quad \forall x \in \Re. \tag{B.34}$$

Especially, if μ is a continuous function, then

$$\mathrm{Cr}\{\xi = x\} = \frac{\mu(x)}{2}, \quad \forall x \in \Re. \tag{B.35}$$

Theorem B.15 *(Sufficient and Necessary Condition for Membership Function) A function $\mu : \Re \to [0,1]$ is a membership function if and only if $\sup \mu(x) = 1$.*

Proof: If μ is a membership function, then there exists a fuzzy variable ξ whose membership function is just μ, and

$$\sup_{x \in \Re} \mu(x) = \sup_{x \in \Re}\left(2\mathrm{Cr}\{\xi = x\}\right) \wedge 1.$$

If there is some point $x \in \Re$ such that $\mathrm{Cr}\{\xi = x\} \geq 0.5$, then $\sup \mu(x) = 1$. Otherwise, we have $\mathrm{Cr}\{\xi = x\} < 0.5$ for each $x \in \Re$. It follows from Axiom 4 that

$$\sup_{x \in \Re} \mu(x) = \sup_{x \in \Re}\left(2\mathrm{Cr}\{\xi = x\}\right) \wedge 1 = 2\sup_{x \in \Re} \mathrm{Cr}\{\xi = x\} = 2\left(\mathrm{Cr}\{\Theta\} \wedge 0.5\right) = 1.$$

Conversely, suppose that $\sup \mu(x) = 1$. For each $x \in \Re$, we define

$$\mathrm{Cr}\{x\} = \frac{1}{2}\left(\mu(x) + 1 - \sup_{y \neq x} \mu(y)\right).$$

It is clear that

$$\sup_{x \in \Re} \mathrm{Cr}\{x\} \geq \frac{1}{2}(1 + 1 - 1) = 0.5.$$

For any $x^* \in \Re$ with $\mathrm{Cr}\{x^*\} \geq 0.5$, we have $\mu(x^*) = 1$ and

$$\mathrm{Cr}\{x^*\} + \sup_{y \neq x^*} \mathrm{Cr}\{y\}$$

$$= \frac{1}{2}\left(\mu(x^*) + 1 - \sup_{y \neq x^*} \mu(y)\right) + \sup_{y \neq x^*} \frac{1}{2}\left(\mu(y) + 1 - \sup_{z \neq y} \mu(z)\right)$$

$$= 1 - \frac{1}{2}\sup_{y \neq x^*}\mu(y) + \frac{1}{2}\sup_{y \neq x^*}\mu(y) = 1.$$

Thus $\mathrm{Cr}\{x\}$ satisfies the credibility extension condition, and has a unique extension to credibility measure on $\mathcal{P}(\Re)$ by using the credibility extension theorem. Now we define a fuzzy variable ξ as an identity function from the credibility space $(\Re, \mathcal{P}(\Re), \mathrm{Cr})$ to $\Re$. Then the membership function of the fuzzy variable ξ is

$$(2\mathrm{Cr}\{\xi = x\}) \wedge 1 = \left(\mu(x) + 1 - \sup_{y \neq x} \mu(y)\right) \wedge 1 = \mu(x)$$

for each x. The theorem is proved.

Some Special Membership Functions

By an *equipossible fuzzy variable* we mean the fuzzy variable fully determined by the pair (a, b) of crisp numbers with $a < b$, whose membership function is given by

$$\mu_1(x) = 1, \quad a \leq x \leq b.$$

By a *triangular fuzzy variable* we mean the fuzzy variable fully determined by the triplet (a, b, c) of crisp numbers with $a < b < c$, whose membership function is given by

$$\mu_2(x) = \begin{cases} \dfrac{x - a}{b - a}, & \text{if } a \leq x \leq b \\[2mm] \dfrac{x - c}{b - c}, & \text{if } b \leq x \leq c. \end{cases}$$

By a *trapezoidal fuzzy variable* we mean the fuzzy variable fully determined by the quadruplet (a, b, c, d) of crisp numbers with $a < b < c < d$, whose membership function is given by

$$\mu_3(x) = \begin{cases} \dfrac{x - a}{b - a}, & \text{if } a \leq x \leq b \\[2mm] 1, & \text{if } b \leq x \leq c \\[2mm] \dfrac{x - d}{c - d}, & \text{if } c \leq x \leq d. \end{cases}$$
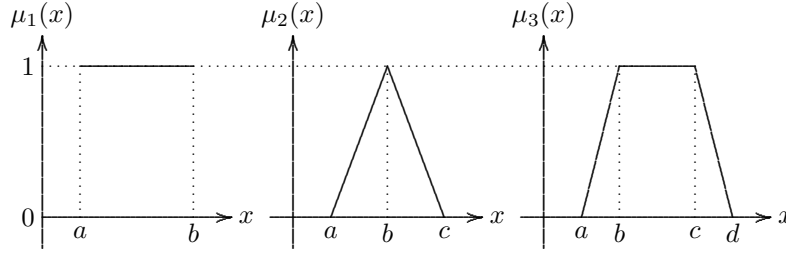
Figure B.1: Membership Functions μ_1, μ_2 and μ_3

B.4   Credibility Distribution

Definition B.9 *(Liu [119]) The credibility distribution $\Phi : \Re \to [0, 1]$ of a fuzzy variable ξ is defined by*

$$\Phi(x) = \text{Cr} \left\{ \theta \in \Theta \mid \xi(\theta) \leq x \right\}. \tag{B.36}$$

That is, $\Phi(x)$ is the credibility that the fuzzy variable ξ takes a value less than or equal to x. Generally speaking, the credibility distribution Φ is neither left-continuous nor right-continuous.

Example B.13: The credibility distribution of an equipossible fuzzy variable (a, b) is

$$\Phi_1(x) = \begin{cases} 0, & \text{if } x < a \\ 1/2, & \text{if } a \leq x < b \\ 1, & \text{if } x \geq b. \end{cases}$$

Especially, if ξ is an equipossible fuzzy variable on $\Re$, then $\Phi_1(x) \equiv 1/2$.

Example B.14: The credibility distribution of a triangular fuzzy variable (a, b, c) is

$$\Phi_2(x) = \begin{cases} 0, & \text{if } x \leq a \\ \dfrac{x - a}{2(b - a)}, & \text{if } a \leq x \leq b \\ \dfrac{x + c - 2b}{2(c - b)}, & \text{if } b \leq x \leq c \\ 1, & \text{if } x \geq c. \end{cases}$$

Example B.15: The credibility distribution of a trapezoidal fuzzy variable (a, b, c, d) is

$$\Phi_3(x) = \begin{cases} 0, & \text{if } x \leq a \\ \dfrac{x-a}{2(b-a)}, & \text{if } a \leq x \leq b \\ \dfrac{1}{2}, & \text{if } b \leq x \leq c \\ \dfrac{x+d-2c}{2(d-c)}, & \text{if } c \leq x \leq d \\ 1, & \text{if } x \geq d. \end{cases}$$
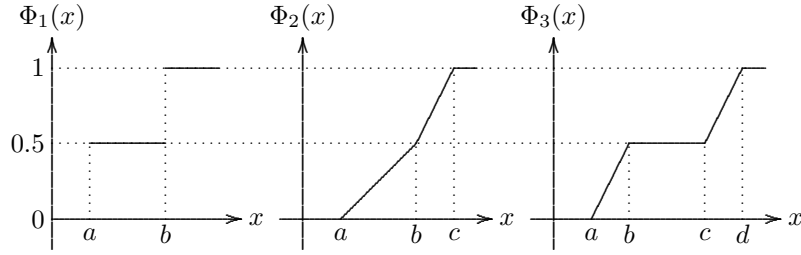
Figure B.2: Credibility Distributions Φ_1, Φ_2 and Φ_3

Theorem B.16 *Let ξ be a fuzzy variable with membership function μ. Then its credibility distribution is*

$$\Phi(x) = \frac{1}{2} \left(\sup_{y \leq x} \mu(y) + 1 - \sup_{y > x} \mu(y) \right), \ \forall x \in \Re. \tag{B.37}$$

Proof: It follows from the credibility inversion theorem immediately.

Theorem B.17 *(Liu [124], Sufficient and Necessary Condition for Credibility Distribution) A function $\Phi : \Re \to [0, 1]$ is a credibility distribution if and only if it is an increasing function with*

$$\lim_{x \to -\infty} \Phi(x) \leq 0.5 \leq \lim_{x \to \infty} \Phi(x), \tag{B.38}$$

$$\lim_{y \downarrow x} \Phi(y) = \Phi(x) \ if \ \lim_{y \downarrow x} \Phi(y) > 0.5 \ or \ \Phi(x) \geq 0.5. \tag{B.39}$$

Proof: It is obvious that a credibility distribution Φ is an increasing function. The inequalities (B.38) follow from the credibility asymptotic theorem immediately. Assume that x is a point at which $\lim_{y \downarrow x} \Phi(y) > 0.5$. That is,

$$\lim_{y \downarrow x} \mathrm{Cr}\{\xi \leq y\} > 0.5.$$

Since $\{\xi \leq y\} \downarrow \{\xi \leq x\}$ as $y \downarrow x$, it follows from the credibility semicontinuity law that

$$\Phi(y) = \mathrm{Cr}\{\xi \leq y\} \downarrow \mathrm{Cr}\{\xi \leq x\} = \Phi(x)$$

as $y \downarrow x$. When x is a point at which $\Phi(x) \geq 0.5$, if $\lim_{y\downarrow x} \Phi(y) \neq \Phi(x)$, then we have

$$\lim_{y\downarrow x} \Phi(y) > \Phi(x) \geq 0.5.$$

For this case, we have proved that $\lim_{y\downarrow x} \Phi(y) = \Phi(x)$. Thus (B.38) and (B.39) are proved. Conversely, if $\Phi : \Re \to [0,1]$ is an increasing function satisfying (B.38) and (B.39), then

$$\mu(x) = \begin{cases} 2\Phi(x), & \text{if } \Phi(x) < 0.5 \\[2mm] 1, & \text{if } \lim_{y\uparrow x}\Phi(y) < 0.5 \leq \Phi(x) \\[2mm] 2 - 2\Phi(x), & \text{if } 0.5 \leq \lim_{y\uparrow x}\Phi(y) \end{cases} \qquad \text{(B.40)}$$

takes values in $[0,1]$ and $\sup \mu(x) = 1$. It follows from Theorem B.15 that there is a fuzzy variable ξ whose membership function is just μ. Let us verify that Φ is the credibility distribution of ξ, i.e., $\mathrm{Cr}\{\xi \leq x\} = \Phi(x)$ for each x. The argument breaks down into two cases. (i) If $\Phi(x) < 0.5$, then we have $\sup_{y>x} \mu(y) = 1$, and $\mu(y) = 2\Phi(y)$ for each y with $y \leq x$. Thus

$$\mathrm{Cr}\{\xi \leq x\} = \frac{1}{2}\left(\sup_{y\leq x}\mu(y) + 1 - \sup_{y>x}\mu(y)\right) = \sup_{y\leq x}\Phi(y) = \Phi(x).$$

(ii) If $\Phi(x) \geq 0.5$, then we have $\sup_{y\leq x} \mu(y) = 1$ and $\Phi(y) \geq \Phi(x) \geq 0.5$ for each y with $y > x$. Thus $\mu(y) = 2 - 2\Phi(y)$ and

$$\begin{aligned}
\mathrm{Cr}\{\xi \leq x\} &= \frac{1}{2}\left(\sup_{y\leq x}\mu(y) + 1 - \sup_{y>x}\mu(y)\right) \\
&= \frac{1}{2}\left(1 + 1 - \sup_{y>x}(2 - 2\Phi(y))\right) \\
&= \inf_{y>x}\Phi(y) = \lim_{y\downarrow x}\Phi(y) = \Phi(x).
\end{aligned}$$

The theorem is proved.

Example B.16: Let a and b be two numbers with $0 \leq a \leq 0.5 \leq b \leq 1$. We define a fuzzy variable by the following membership function,

$$\mu(x) = \begin{cases} 2a, & \text{if } x < 0 \\ 1, & \text{if } x = 0 \\ 2 - 2b, & \text{if } x > 0. \end{cases}$$

Then its credibility distribution is

$$\Phi(x) = \begin{cases} a, & \text{if } x < 0 \\ b, & \text{if } x \geq 0. \end{cases}$$

Thus we have

$$\lim_{x \to -\infty} \Phi(x) = a, \qquad \lim_{x \to +\infty} \Phi(x) = b.$$

B.5   Independence

The independence of fuzzy variables has been discussed by many authors from different angles. Here we use the following definition.

Definition B.10 *(Liu and Gao [150]) The fuzzy variables $\xi_1, \xi_2, \cdots, \xi_m$ are said to be independent if*

$$\text{Cr} \left\{ \bigcap_{i=1}^{m} \{\xi_i \in B_i\} \right\} = \min_{1 \leq i \leq m} \text{Cr} \left\{ \xi_i \in B_i \right\} \tag{B.41}$$

for any sets $B_1, B_2, \cdots, B_m$ of $\Re$.

Theorem B.18 *The fuzzy variables $\xi_1, \xi_2, \cdots, \xi_m$ are independent if and only if*

$$\text{Cr} \left\{ \bigcup_{i=1}^{m} \{\xi_i \in B_i\} \right\} = \max_{1 \leq i \leq m} \text{Cr} \left\{ \xi_i \in B_i \right\} \tag{B.42}$$

for any sets $B_1, B_2, \cdots, B_m$ of $\Re$.

Proof: It follows from the self-duality of credibility measure that $\xi_1, \xi_2, \cdots, \xi_m$ are independent if and only if

$$\text{Cr} \left\{ \bigcup_{i=1}^{m} \{\xi_i \in B_i\} \right\} = 1 - \text{Cr} \left\{ \bigcap_{i=1}^{m} \{\xi_i \in B_i^c\} \right\}$$

$$= 1 - \min_{1 \leq i \leq m} \text{Cr}\{\xi_i \in B_i^c\} = \max_{1 \leq i \leq m} \text{Cr} \left\{ \xi_i \in B_i \right\}.$$

Thus (B.42) is verified. The proof is complete.

Theorem B.19 *The fuzzy variables $\xi_1, \xi_2, \cdots, \xi_m$ are independent if and only if*

$$\text{Cr} \left\{ \bigcap_{i=1}^{m} \{\xi_i = x_i\} \right\} = \min_{1 \leq i \leq m} \text{Cr} \left\{ \xi_i = x_i \right\} \tag{B.43}$$

for any real numbers $x_1, x_2, \cdots, x_m$.

Proof: If $\xi_1, \xi_2, \cdots, \xi_m$ are independent, then we have (B.43) immediately by taking $B_i = \{x_i\}$ for each i. Conversely, if $\mathrm{Cr}\{\cap_{i=1}^m (\xi_i \in B_i)\} \geq 0.5$, it follows from Theorem B.2 that (B.41) holds. Otherwise, we have $\mathrm{Cr}\{\cap_{i=1}^m (\xi_i = x_i)\} < 0.5$ for any real numbers $x_i \in B_i$, $i = 1, 2, \cdots, m$, and

$$\mathrm{Cr}\left\{\bigcap_{i=1}^m \{\xi_i \in B_i\}\right\} = \mathrm{Cr}\left\{\bigcup_{x_i \in B_i, 1 \leq i \leq m} \bigcap_{i=1}^m \{\xi_i = x_i\}\right\}$$

$$= \sup_{x_i \in B_i, 1 \leq i \leq m} \mathrm{Cr}\left\{\bigcap_{i=1}^m \{\xi_i = x_i\}\right\} = \sup_{x_i \in B_i, 1 \leq i \leq m} \min_{1 \leq i \leq m} \mathrm{Cr}\{\xi_i = x_i\}$$

$$= \min_{1 \leq i \leq m} \sup_{x_i \in B_i} \mathrm{Cr}\{\xi_i = x_i\} = \min_{1 \leq i \leq m} \mathrm{Cr}\{\xi_i \in B_i\}.$$

Hence (B.41) is true, and $\xi_1, \xi_2, \cdots, \xi_m$ are independent. The theorem is thus proved.

Theorem B.20 *Let μ_i be membership functions of fuzzy variables ξ_i, $i = 1, 2, \cdots, m$, respectively, and μ the joint membership function of fuzzy vector $(\xi_1, \xi_2, \cdots, \xi_m)$. Then the fuzzy variables $\xi_1, \xi_2, \cdots, \xi_m$ are independent if and only if*

$$\mu(x_1, x_2, \cdots, x_m) = \min_{1 \leq i \leq m} \mu_i(x_i) \tag{B.44}$$

for any real numbers $x_1, x_2, \cdots, x_m$.

Proof: Suppose that $\xi_1, \xi_2, \cdots, \xi_m$ are independent. It follows from Theorem B.19 that

$$\mu(x_1, x_2, \cdots, x_m) = \left(2\mathrm{Cr}\left\{\bigcap_{i=1}^m \{\xi_i = x_i\}\right\}\right) \wedge 1$$

$$= \left(2 \min_{1 \leq i \leq m} \mathrm{Cr}\{\xi_i = x_i\}\right) \wedge 1$$

$$= \min_{1 \leq i \leq m} \left(2\mathrm{Cr}\{\xi_i = x_i\}\right) \wedge 1 = \min_{1 \leq i \leq m} \mu_i(x_i).$$

Conversely, for any real numbers $x_1, x_2, \cdots, x_m$ with $\mathrm{Cr}\{\cap_{i=1}^m \{\xi_i = x_i\}\} < 0.5$, we have

$$\mathrm{Cr}\left\{\bigcap_{i=1}^m \{\xi_i = x_i\}\right\} = \frac{1}{2}\left(2\mathrm{Cr}\left\{\bigcap_{i=1}^m \{\xi_i = x_i\}\right\}\right) \wedge 1$$

$$= \frac{1}{2}\mu(x_1, x_2, \cdots, x_m) = \frac{1}{2} \min_{1 \leq i \leq m} \mu_i(x_i)$$

$$= \frac{1}{2}\left(\min_{1 \leq i \leq m} \left(2\mathrm{Cr}\{\xi_i = x_i\}\right) \wedge 1\right)$$

$$= \min_{1 \leq i \leq m} \mathrm{Cr}\{\xi_i = x_i\}.$$

It follows from Theorem B.19 that $\xi_1, \xi_2, \cdots, \xi_m$ are independent. The theorem is proved.

B.6   Extension Principle of Zadeh

Theorem B.21 *(Extension Principle of Zadeh) Let $\xi_1, \xi_2, \cdots, \xi_n$ be independent fuzzy variables with membership functions $\mu_1, \mu_2, \cdots, \mu_n$, respectively, and $f : \Re^n \rightarrow \Re$ a function. Then the membership function μ of $\xi = f(\xi_1, \xi_2, \cdots, \xi_n)$ is derived from the membership functions $\mu_1, \mu_2, \cdots, \mu_n$ by*

$$\mu(x) = \sup_{x=f(x_1,x_2,\cdots,x_n)} \min_{1 \leq i \leq n} \mu_i(x_i) \qquad (B.45)$$

for any $x \in \Re$. Here we set $\mu(x) = 0$ if there are not real numbers $x_1, x_2, \cdots, x_n$ such that $x = f(x_1, x_2, \cdots, x_n)$.

Proof: It follows from Definition B.8 that the membership function of $\xi = f(\xi_1, \xi_2, \cdots, \xi_n)$ is

$$\mu(x) = (2\mathrm{Cr}\,\{f(\xi_1, \xi_2, \cdots, \xi_n) = x\}) \wedge 1$$

$$= \left(2 \sup_{x=f(x_1,x_2,\cdots,x_n)} \mathrm{Cr}\{\xi_1 = x_1, \xi_2 = x_2, \cdots, \xi_n = x_n\} \right) \wedge 1$$

$$= \left(2 \sup_{x=f(x_1,x_2,\cdots,x_n)} \min_{1 \leq k \leq n} \mathrm{Cr}\{\xi_i = x_i\} \right) \wedge 1 \quad \text{(by independence)}$$

$$= \sup_{x=f(x_1,x_2,\cdots,x_n)} \min_{1 \leq k \leq n} (2\mathrm{Cr}\{\xi_i = x_i\}) \wedge 1$$

$$= \sup_{x=f(x_1,x_2,\cdots,x_n)} \min_{1 \leq i \leq n} \mu_i(x_i).$$

The theorem is proved.

Remark B.3: The extension principle of Zadeh is only applicable to the operations on independent fuzzy variables. In the past literature, the extension principle is used as a postulate. However, it is treated as a theorem in credibility theory.

Example B.17: The sum of independent equipossible fuzzy variables $\xi = (a_1, a_2)$ and $\eta = (b_1, b_2)$ is also an equipossible fuzzy variable, and

$$\xi + \eta = (a_1 + b_1, a_2 + b_2).$$

Their product is also an equipossible fuzzy variable, and

$$\xi \cdot \eta = \left(\min_{a_1 \leq x \leq a_2, b_1 \leq y \leq b_2} xy, \ \max_{a_1 \leq x \leq a_2, b_1 \leq y \leq b_2} xy \right).$$

Example B.18: The sum of independent triangular fuzzy variables $\xi = (a_1, a_2, a_3)$ and $\eta = (b_1, b_2, b_3)$ is also a triangular fuzzy variable, and

$$\xi + \eta = (a_1 + b_1, a_2 + b_2, a_3 + b_3).$$

The product of a triangular fuzzy variable $\xi = (a_1, a_2, a_3)$ and a scalar number λ is

$$\lambda \cdot \xi = \begin{cases} (\lambda a_1, \lambda a_2, \lambda a_3), & \text{if } \lambda \geq 0 \\ (\lambda a_3, \lambda a_2, \lambda a_1), & \text{if } \lambda < 0. \end{cases}$$

That is, the product of a triangular fuzzy variable and a scalar number is also a triangular fuzzy variable. However, the product of two triangular fuzzy variables is not a triangular one.

Example B.19: The sum of independent trapezoidal fuzzy variables $\xi = (a_1, a_2, a_3, a_4)$ and $\eta = (b_1, b_2, b_3, b_4)$ is also a trapezoidal fuzzy variable, and

$$\xi + \eta = (a_1 + b_1, a_2 + b_2, a_3 + b_3, a_4 + b_4).$$

The product of a trapezoidal fuzzy variable $\xi = (a_1, a_2, a_3, a_4)$ and a scalar number λ is

$$\lambda \cdot \xi = \begin{cases} (\lambda a_1, \lambda a_2, \lambda a_3, \lambda a_4), & \text{if } \lambda \geq 0 \\ (\lambda a_4, \lambda a_3, \lambda a_2, \lambda a_1), & \text{if } \lambda < 0. \end{cases}$$

That is, the product of a trapezoidal fuzzy variable and a scalar number is also a trapezoidal fuzzy variable. However, the product of two trapezoidal fuzzy variables is not a trapezoidal one.

Example B.20: Let $\xi_1, \xi_2, \cdots, \xi_n$ be independent fuzzy variables with membership functions $\mu_1, \mu_2, \cdots, \mu_n$, respectively, and $f : \Re^n \to \Re$ a function. Then for any set B of real numbers, the credibility $\text{Cr}\{f(\xi_1, \xi_2, \cdots, \xi_n) \in B\}$ is

$$\frac{1}{2} \left(\sup_{f(x_1, x_2, \cdots, x_n) \in B} \min_{1 \leq i \leq n} \mu_i(x_i) + 1 - \sup_{f(x_1, x_2, \cdots, x_n) \in B^c} \min_{1 \leq i \leq n} \mu_i(x_i) \right).$$

B.7   Expected Value

There are many ways to define an expected value operator for fuzzy variables. The most general definition of expected value operator of fuzzy variable was given by Liu and Liu [121]. This definition is applicable to not only continuous fuzzy variables but also discrete ones.

Definition B.11 *(Liu and Liu [121]) Let ξ be a fuzzy variable. Then the expected value of ξ is defined by*

$$E[\xi] = \int_0^{+\infty} \text{Cr}\{\xi \geq r\} \mathrm{d}r - \int_{-\infty}^0 \text{Cr}\{\xi \leq r\} \mathrm{d}r \tag{B.46}$$

provided that at least one of the two integrals is finite.

Example B.21: Let ξ be the equipossible fuzzy variable (a, b). Then its expected value is $E[\xi] = (a + b)/2$.

Example B.22: The triangular fuzzy variable $\xi = (a, b, c)$ has an expected value $E[\xi] = (a + 2b + c)/4$.

Example B.23: The trapezoidal fuzzy variable $\xi = (a, b, c, d)$ has an expected value $E[\xi] = (a + b + c + d)/4$.

Example B.24: Let ξ be a continuous fuzzy variable with membership function μ. If its expected value exists, and there is a point x_0 such that $\mu(x)$ is increasing on $(-\infty, x_0)$ and decreasing on $(x_0, +\infty)$, then $\mathrm{Cr}\{\xi \geq x\} = \mu(x)/2$ for any $x > x_0$ and $\mathrm{Cr}\{\xi \leq x\} = \mu(x)/2$ for any $x < x_0$. Thus

$$E[\xi] = x_0 + \frac{1}{2} \int_{x_0}^{+\infty} \mu(x)\mathrm{d}x - \frac{1}{2} \int_{-\infty}^{x_0} \mu(x)\mathrm{d}x.$$

Example B.25: The definition of expected value operator is also applicable to discrete case. Assume that ξ is a simple fuzzy variable whose membership function is given by

$$\mu(x) = \begin{cases} \mu_1, & \text{if } x = x_1 \\ \mu_2, & \text{if } x = x_2 \\ \cdots \\ \mu_m, & \text{if } x = x_m \end{cases} \tag{B.47}$$

where $x_1, x_2, \cdots, x_m$ are distinct numbers. Note that $\mu_1 \vee \mu_2 \vee \cdots \vee \mu_m = 1$. Definition B.11 implies that the expected value of ξ is

$$E[\xi] = \sum_{i=1}^{m} w_i x_i \tag{B.48}$$

where the weights are given by

$$w_i = \frac{1}{2} \left(\max_{1 \leq j \leq m} \{\mu_j | x_j \leq x_i\} - \max_{1 \leq j \leq m} \{\mu_j | x_j < x_i\} \right.$$
$$\left. + \max_{1 \leq j \leq m} \{\mu_j | x_j \geq x_i\} - \max_{1 \leq j \leq m} \{\mu_j | x_j > x_i\} \right)$$

for $i = 1, 2, \cdots, m$. It is easy to verify that all $w_i \geq 0$ and the sum of all weights is just 1.

Example B.26: Consider the fuzzy variable ξ defined by (B.47). Suppose $x_1 < x_2 < \cdots < x_m$ and there exists an index k with $1 < k < m$ such that

$$\mu_1 \leq \mu_2 \leq \cdots \leq \mu_k \quad \text{and} \quad \mu_k \geq \mu_{k+1} \geq \cdots \geq \mu_m.$$

Note that $\mu_k \equiv 1$. Then the expected value is determined by (B.48) and the weights are given by

$$
w_i = \begin{cases}
\dfrac{\mu_1}{2}, & \text{if } i = 1 \\[2mm]
\dfrac{\mu_i - \mu_{i-1}}{2}, & \text{if } i = 2, 3, \cdots, k - 1 \\[2mm]
1 - \dfrac{\mu_{k-1} + \mu_{k+1}}{2}, & \text{if } i = k \\[2mm]
\dfrac{\mu_i - \mu_{i+1}}{2}, & \text{if } i = k + 1, k + 2, \cdots, m - 1 \\[2mm]
\dfrac{\mu_m}{2}, & \text{if } i = m.
\end{cases}
$$

Linearity of Expected Value Operator

Theorem B.22 *(Liu and Liu [143]) Let ξ and η be independent fuzzy variables with finite expected values. Then for any numbers a and b, we have*

$$
E[a\xi + b\eta] = aE[\xi] + bE[\eta]. \tag{B.49}
$$

Proof: STEP 1: We first prove that $E[\xi + b] = E[\xi] + b$ for any real number b. If $b \geq 0$, we have

$$
\begin{aligned}
E[\xi + b] &= \int_0^\infty \mathrm{Cr}\{\xi + b \geq r\}\mathrm{d}r - \int_{-\infty}^0 \mathrm{Cr}\{\xi + b \leq r\}\mathrm{d}r \\
&= \int_0^\infty \mathrm{Cr}\{\xi \geq r - b\}\mathrm{d}r - \int_{-\infty}^0 \mathrm{Cr}\{\xi \leq r - b\}\mathrm{d}r \\
&= E[\xi] + \int_0^b \left(\mathrm{Cr}\{\xi \geq r - b\} + \mathrm{Cr}\{\xi < r - b\} \right) \mathrm{d}r \\
&= E[\xi] + b.
\end{aligned}
$$

If $b < 0$, then we have

$$
E[\xi + b] = E[\xi] - \int_b^0 \left(\mathrm{Cr}\{\xi \geq r - b\} + \mathrm{Cr}\{\xi < r - b\} \right) \mathrm{d}r = E[\xi] + b.
$$

STEP 2: We prove that $E[a\xi] = aE[\xi]$ for any real number a. If $a = 0$, then the equation $E[a\xi] = aE[\xi]$ holds trivially. If $a > 0$, we have

$$
\begin{aligned}
E[a\xi] &= \int_0^\infty \mathrm{Cr}\{a\xi \geq r\}\mathrm{d}r - \int_{-\infty}^0 \mathrm{Cr}\{a\xi \leq r\}\mathrm{d}r \\
&= \int_0^\infty \mathrm{Cr}\left\{\xi \geq \frac{r}{a}\right\}\mathrm{d}r - \int_{-\infty}^0 \mathrm{Cr}\left\{\xi \leq \frac{r}{a}\right\}\mathrm{d}r \\
&= a\int_0^\infty \mathrm{Cr}\left\{\xi \geq \frac{r}{a}\right\}\mathrm{d}\left(\frac{r}{a}\right) - a\int_{-\infty}^0 \mathrm{Cr}\left\{\xi \leq \frac{r}{a}\right\}\mathrm{d}\left(\frac{r}{a}\right) = aE[\xi].
\end{aligned}
$$

If $a < 0$, we have

$$
\begin{aligned}
E[a\xi] &= \int_0^\infty \mathrm{Cr}\{a\xi \geq r\}\mathrm{d}r - \int_{-\infty}^0 \mathrm{Cr}\{a\xi \leq r\}\mathrm{d}r \\
&= \int_0^\infty \mathrm{Cr}\left\{\xi \leq \frac{r}{a}\right\}\mathrm{d}r - \int_{-\infty}^0 \mathrm{Cr}\left\{\xi \geq \frac{r}{a}\right\}\mathrm{d}r \\
&= a\int_0^\infty \mathrm{Cr}\left\{\xi \geq \frac{r}{a}\right\}\mathrm{d}\left(\frac{r}{a}\right) - a\int_{-\infty}^0 \mathrm{Cr}\left\{\xi \leq \frac{r}{a}\right\}\mathrm{d}\left(\frac{r}{a}\right) = aE[\xi].
\end{aligned}
$$

STEP 3: We prove that $E[\xi + \eta] = E[\xi] + E[\eta]$ when both ξ and η are simple fuzzy variables with the following membership functions,

$$
\mu(x) = \begin{cases} \mu_1, & \text{if } x = a_1 \\ \mu_2, & \text{if } x = a_2 \\ \cdots \\ \mu_m, & \text{if } x = a_m, \end{cases} \qquad \nu(x) = \begin{cases} \nu_1, & \text{if } x = b_1 \\ \nu_2, & \text{if } x = b_2 \\ \cdots \\ \nu_n, & \text{if } x = b_n. \end{cases}
$$

Then $\xi + \eta$ is also a simple fuzzy variable taking values $a_i + b_j$ with membership degrees $\mu_i \wedge \nu_j$, $i = 1, 2, \cdots, m$, $j = 1, 2, \cdots, n$, respectively. Now we define

$$
\begin{aligned}
w_i' = \frac{1}{2}\Bigg(&\max_{1 \leq k \leq m}\{\mu_k | a_k \leq a_i\} - \max_{1 \leq k \leq m}\{\mu_k | a_k < a_i\} \\
&+ \max_{1 \leq k \leq m}\{\mu_k | a_k \geq a_i\} - \max_{1 \leq k \leq m}\{\mu_k | a_k > a_i\}\Bigg),
\end{aligned}
$$

$$
\begin{aligned}
w_j'' = \frac{1}{2}\Bigg(&\max_{1 \leq l \leq n}\{\nu_l | b_l \leq b_j\} - \max_{1 \leq l \leq n}\{\nu_l | b_l < b_j\} \\
&+ \max_{1 \leq l \leq n}\{\nu_l | b_l \geq b_j\} - \max_{1 \leq l \leq n}\{\nu_l | b_l > b_j\}\Bigg),
\end{aligned}
$$

$$
\begin{aligned}
w_{ij} = \frac{1}{2}\Bigg(&\max_{1 \leq k \leq m, 1 \leq l \leq n}\{\mu_k \wedge \nu_l | a_k + b_l \leq a_i + b_j\} \\
&- \max_{1 \leq k \leq m, 1 \leq l \leq n}\{\mu_k \wedge \nu_l | a_k + b_l < a_i + b_j\} \\
&+ \max_{1 \leq k \leq m, 1 \leq l \leq n}\{\mu_k \wedge \nu_l | a_k + b_l \geq a_i + b_j\} \\
&- \max_{1 \leq k \leq m, 1 \leq l \leq n}\{\mu_k \wedge \nu_l | a_k + b_l > a_i + b_j\}\Bigg)
\end{aligned}
$$

for $i = 1, 2, \cdots, m$ and $j = 1, 2, \cdots, n$. It is also easy to verify that

$$
w_i' = \sum_{j=1}^n w_{ij}, \qquad w_j'' = \sum_{i=1}^m w_{ij}
$$

for $i = 1, 2, \cdots, m$ and $j = 1, 2, \cdots, n$. If $\{a_i\}$, $\{b_j\}$ and $\{a_i + b_j\}$ are sequences consisting of distinct elements, then

$$E[\xi] = \sum_{i=1}^{m} a_i w_i', \quad E[\eta] = \sum_{j=1}^{n} b_j w_j'', \quad E[\xi + \eta] = \sum_{i=1}^{m} \sum_{j=1}^{n} (a_i + b_j) w_{ij}.$$

Thus $E[\xi + \eta] = E[\xi] + E[\eta]$. If not, we may give them a small perturbation such that they are distinct, and prove the linearity by letting the perturbation tend to zero.

STEP 4: We prove that $E[\xi + \eta] = E[\xi] + E[\eta]$ when ξ and η are fuzzy variables such that

$$
\begin{aligned}
\lim_{y \uparrow 0} \text{Cr}\{\xi \le y\} \le \frac{1}{2} \le \text{Cr}\{\xi \le 0\}, \\
\lim_{y \uparrow 0} \text{Cr}\{\eta \le y\} \le \frac{1}{2} \le \text{Cr}\{\eta \le 0\}.
\end{aligned}
\tag{B.50}
$$

We define simple fuzzy variables ξ_i via credibility distributions as follows,

$$
\Phi_i(x) = \begin{cases}
\dfrac{k-1}{2^i}, & \text{if } \dfrac{k-1}{2^i} \le \text{Cr}\{\xi \le x\} < \dfrac{k}{2^i}, \ k = 1, 2, \cdots, 2^{i-1} \\[2ex]
\dfrac{k}{2^i}, & \text{if } \dfrac{k-1}{2^i} \le \text{Cr}\{\xi \le x\} < \dfrac{k}{2^i}, \ k = 2^{i-1} + 1, \cdots, 2^i \\[2ex]
1, & \text{if } \text{Cr}\{\xi \le x\} = 1
\end{cases}
$$

for $i = 1, 2, \cdots$ Thus $\{\xi_i\}$ is a sequence of simple fuzzy variables satisfying

$$
\begin{aligned}
\text{Cr}\{\xi_i \le r\} \uparrow \text{Cr}\{\xi \le r\}, & \quad \text{if } r \le 0 \\
\text{Cr}\{\xi_i \ge r\} \uparrow \text{Cr}\{\xi \ge r\}, & \quad \text{if } r \ge 0
\end{aligned}
$$

as $i \to \infty$. Similarly, we define simple fuzzy variables η_i via credibility distributions as follows,

$$
\Psi_i(x) = \begin{cases}
\dfrac{k-1}{2^i}, & \text{if } \dfrac{k-1}{2^i} \le \text{Cr}\{\eta \le x\} < \dfrac{k}{2^i}, \ k = 1, 2, \cdots, 2^{i-1} \\[2ex]
\dfrac{k}{2^i}, & \text{if } \dfrac{k-1}{2^i} \le \text{Cr}\{\eta \le x\} < \dfrac{k}{2^i}, \ k = 2^{i-1} + 1, \cdots, 2^i \\[2ex]
1, & \text{if } \text{Cr}\{\eta \le x\} = 1
\end{cases}
$$

for $i = 1, 2, \cdots$ Thus $\{\eta_i\}$ is a sequence of simple fuzzy variables satisfying

$$
\begin{aligned}
\text{Cr}\{\eta_i \le r\} \uparrow \text{Cr}\{\eta \le r\}, & \quad \text{if } r \le 0 \\
\text{Cr}\{\eta_i \ge r\} \uparrow \text{Cr}\{\eta \ge r\}, & \quad \text{if } r \ge 0
\end{aligned}
$$

as $i \to \infty$. It is also clear that $\{\xi_i + \eta_i\}$ is a sequence of simple fuzzy variables. Furthermore, when $r \leq 0$, it follows from (B.50) that

$$
\begin{aligned}
\lim_{i\to\infty} \mathrm{Cr}\{\xi_i + \eta_i \leq r\} &= \lim_{i\to\infty} \sup_{x\leq 0, y\leq 0, x+y\leq r} \mathrm{Cr}\{\xi_i \leq x\} \wedge \mathrm{Cr}\{\eta_i \leq y\} \\
&= \sup_{x\leq 0, y\leq 0, x+y\leq r} \lim_{i\to\infty} \mathrm{Cr}\{\xi_i \leq x\} \wedge \mathrm{Cr}\{\eta_i \leq y\} \\
&= \sup_{x\leq 0, y\leq 0, x+y\leq r} \mathrm{Cr}\{\xi \leq x\} \wedge \mathrm{Cr}\{\eta \leq y\} \\
&= \mathrm{Cr}\{\xi + \eta \leq r\}.
\end{aligned}
$$

That is,
$$
\mathrm{Cr}\{\xi_i + \eta_i \leq r\} \uparrow \mathrm{Cr}\{\xi + \eta \leq r\}, \quad \text{if } r \leq 0.
$$

A similar way may prove that

$$
\mathrm{Cr}\{\xi_i + \eta_i \geq r\} \uparrow \mathrm{Cr}\{\xi + \eta \geq r\}, \quad \text{if } r \geq 0.
$$

Since the expected values $E[\xi]$ and $E[\eta]$ exist, we have

$$
\begin{aligned}
E[\xi_i] &= \int_0^{+\infty} \mathrm{Cr}\{\xi_i \geq r\} \mathrm{d}r - \int_{-\infty}^0 \mathrm{Cr}\{\xi_i \leq r\} \mathrm{d}r \\
&\to \int_0^{+\infty} \mathrm{Cr}\{\xi \geq r\} \mathrm{d}r - \int_{-\infty}^0 \mathrm{Cr}\{\xi \leq r\} \mathrm{d}r = E[\xi],
\end{aligned}
$$

$$
\begin{aligned}
E[\eta_i] &= \int_0^{+\infty} \mathrm{Cr}\{\eta_i \geq r\} \mathrm{d}r - \int_{-\infty}^0 \mathrm{Cr}\{\eta_i \leq r\} \mathrm{d}r \\
&\to \int_0^{+\infty} \mathrm{Cr}\{\eta \geq r\} \mathrm{d}r - \int_{-\infty}^0 \mathrm{Cr}\{\eta \leq r\} \mathrm{d}r = E[\eta],
\end{aligned}
$$

$$
\begin{aligned}
E[\xi_i + \eta_i] &= \int_0^{+\infty} \mathrm{Cr}\{\xi_i + \eta_i \geq r\} \mathrm{d}r - \int_{-\infty}^0 \mathrm{Cr}\{\xi_i + \eta_i \leq r\} \mathrm{d}r \\
&\to \int_0^{+\infty} \mathrm{Cr}\{\xi + \eta \geq r\} \mathrm{d}r - \int_{-\infty}^0 \mathrm{Cr}\{\xi + \eta \leq r\} \mathrm{d}r = E[\xi + \eta]
\end{aligned}
$$

as $i \to \infty$. It follows from Step 3 that $E[\xi + \eta] = E[\xi] + E[\eta]$.

STEP 5: We prove that $E[\xi + \eta] = E[\xi] + E[\eta]$ when ξ and η are arbitrary fuzzy variables. Since they have finite expected values, there exist two numbers c and d such that

$$
\lim_{y\uparrow 0} \mathrm{Cr}\{\xi + c \leq y\} \leq \frac{1}{2} \leq \mathrm{Cr}\{\xi + c \leq 0\},
$$

$$
\lim_{y\uparrow 0} \mathrm{Cr}\{\eta + d \leq y\} \leq \frac{1}{2} \leq \mathrm{Cr}\{\eta + d \leq 0\}.
$$

It follows from Steps 1 and 4 that

$$\begin{aligned}
E[\xi + \eta] &= E[(\xi + c) + (\eta + d) - c - d] \\
&= E[(\xi + c) + (\eta + d)] - c - d \\
&= E[\xi + c] + E[\eta + d] - c - d \\
&= E[\xi] + c + E[\eta] + d - c - d \\
&= E[\xi] + E[\eta].
\end{aligned}$$

STEP 6: We prove that $E[a\xi + b\eta] = aE[\xi] + bE[\eta]$ for any real numbers a and b. In fact, the equation follows immediately from Steps 2 and 5. The theorem is proved.

B.8   Variance

Definition B.12 *(Liu and Liu [121]) Let ξ be a fuzzy variable with finite expected value e. Then the variance of ξ is defined by $V[\xi] = E[(\xi - e)^2]$.*

The variance of a fuzzy variable provides a measure of the spread of the distribution around its expected value.

Example B.27: A fuzzy variable ξ is called normally distributed if it has a normal membership function

$$\mu(x) = 2\left(1 + \exp\left(\frac{\pi|x - e|}{\sqrt{6}\sigma}\right)\right)^{-1}, \quad x \in \Re, \sigma > 0. \tag{B.51}$$

The expected value is e and variance is σ^2. Let ξ_1 and ξ_2 be independently and normally distributed fuzzy variables with expected values e_1 and e_2, variances σ_1^2 and σ_2^2, respectively. Then for any real numbers a_1 and a_2, the fuzzy variable $a_1\xi_1 + a_2\xi_2$ is also normally distributed with expected value $a_1e_1 + a_2e_2$ and variance $(|a_1|\sigma_1 + |a_2|\sigma_2)^2$.
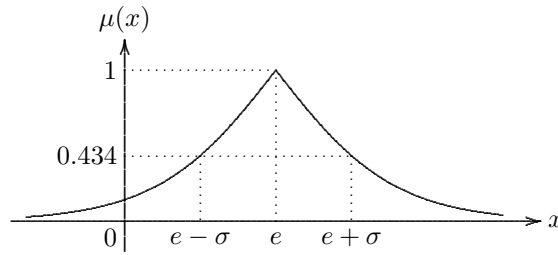
Figure B.3: Normal Membership Function

Theorem B.23 *If ξ is a fuzzy variable whose variance exists, a and b are real numbers, then $V[a\xi + b] = a^2 V[\xi]$.*

Proof: It follows from the definition of variance that

$$V[a\xi + b] = E\left[(a\xi + b - aE[\xi] - b)^2\right] = a^2 E[(\xi - E[\xi])^2] = a^2 V[\xi].$$

Theorem B.24 *Let ξ be a fuzzy variable with expected value e. Then $V[\xi] = 0$ if and only if $\mathrm{Cr}\{\xi = e\} = 1$.*

Proof: If $V[\xi] = 0$, then $E[(\xi - e)^2] = 0$. Note that

$$E[(\xi - e)^2] = \int_0^{+\infty} \mathrm{Cr}\{(\xi - e)^2 \geq r\}\mathrm{d}r$$

which implies $\mathrm{Cr}\{(\xi - e)^2 \geq r\} = 0$ for any $r > 0$. Hence we have $\mathrm{Cr}\{(\xi - e)^2 = 0\} = 1$, i.e., $\mathrm{Cr}\{\xi = e\} = 1$. Conversely, if $\mathrm{Cr}\{\xi = e\} = 1$, then we have $\mathrm{Cr}\{(\xi - e)^2 = 0\} = 1$ and $\mathrm{Cr}\{(\xi - e)^2 \geq r\} = 0$ for any $r > 0$. Thus

$$V[\xi] = \int_0^{+\infty} \mathrm{Cr}\{(\xi - e)^2 \geq r\}\mathrm{d}r = 0.$$

B.9   Moments

Definition B.13 *(Liu [123]) Let ξ be a fuzzy variable, and k a positive number. Then*
(a) the expected value $E[\xi^k]$ is called the kth moment;
(b) the expected value $E[|\xi|^k]$ is called the kth absolute moment;
(c) the expected value $E[(\xi - E[\xi])^k]$ is called the kth central moment;
(d) the expected value $E[|\xi - E[\xi]|^k]$ is called the kth absolute central moment.

Note that the first central moment is always 0, the first moment is just the expected value, and the second central moment is just the variance.

Example B.28: A fuzzy variable ξ is called exponentially distributed if it has an exponential membership function

$$\mu(x) = 2\left(1 + \exp\left(\frac{\pi x}{\sqrt{6}m}\right)\right)^{-1}, \quad x \geq 0, m > 0. \tag{B.52}$$

The expected value is $(\sqrt{6}m\ln 2)/\pi$ and the second moment is m^2. Let ξ_1 and ξ_2 be independently and exponentially distributed fuzzy variables with second moments m_1^2 and m_2^2, respectively. Then for any positive real numbers a_1 and a_2, the fuzzy variable $a_1\xi_1 + a_2\xi_2$ is also exponentially distributed with second moment $(a_1 m_1 + a_2 m_2)^2$.
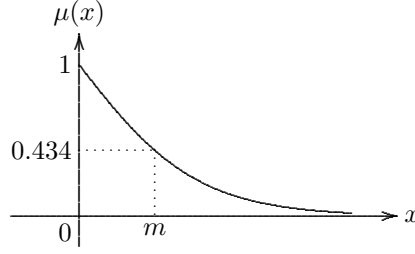
Figure B.4: Exponential Membership Function

B.10   Critical Values

In order to rank fuzzy variables, we may use two critical values: optimistic value and pessimistic value.

Definition B.14 *(Liu [119]) Let ξ be a fuzzy variable, and $\alpha \in (0, 1]$. Then*

$$\xi_{\sup}(\alpha) = \sup \left\{ r \mid \text{Cr} \left\{ \xi \geq r \right\} \geq \alpha \right\} \qquad (B.53)$$

is called the α-optimistic value to ξ, and

$$\xi_{\inf}(\alpha) = \inf \left\{ r \mid \text{Cr} \left\{ \xi \leq r \right\} \geq \alpha \right\} \qquad (B.54)$$

is called the α-pessimistic value to ξ.

Example B.29: Let $\xi = (a, b, c)$ be a triangular fuzzy variable. Then its α-optimistic and α-pessimistic values are

$$\xi_{\sup}(\alpha) = \begin{cases} 2\alpha b + (1 - 2\alpha)c, & \text{if } \alpha \leq 0.5 \\ (2\alpha - 1)a + (2 - 2\alpha)b, & \text{if } \alpha > 0.5, \end{cases}$$

$$\xi_{\inf}(\alpha) = \begin{cases} (1 - 2\alpha)a + 2\alpha b, & \text{if } \alpha \leq 0.5 \\ (2 - 2\alpha)b + (2\alpha - 1)c, & \text{if } \alpha > 0.5. \end{cases}$$

Theorem B.25 *Let ξ be a fuzzy variable, and $\alpha \in (0, 1]$. If $\alpha > 0.5$, then we have*

$$\text{Cr}\{\xi \leq \xi_{\inf}(\alpha)\} \geq \alpha, \quad \text{Cr}\{\xi \geq \xi_{\sup}(\alpha)\} \geq \alpha. \qquad (B.55)$$

Proof: It follows from the definition of α-pessimistic value that there exists a decreasing sequence $\{x_i\}$ such that $\text{Cr}\{\xi \leq x_i\} \geq \alpha$ and $x_i \downarrow \xi_{\inf}(\alpha)$ as $i \to \infty$. Since $\{\xi \leq x_i\} \downarrow \{\xi \leq \xi_{\inf}(\alpha)\}$ and $\lim_{i \to \infty} \text{Cr}\{\xi \leq x_i\} \geq \alpha > 0.5$, it follows from the credibility semicontinuity law that

$$\text{Cr}\{\xi \leq \xi_{\inf}(\alpha)\} = \lim_{i \to \infty} \text{Cr}\{\xi \leq x_i\} \geq \alpha.$$

Similarly, there exists an increasing sequence $\{x_i\}$ such that $\mathrm{Cr}\{\xi \geq x_i\} \geq \alpha$ and $x_i \uparrow \xi_{\sup}(\alpha)$ as $i \to \infty$. Since $\{\xi \geq x_i\} \downarrow \{\xi \geq \xi_{\sup}(\alpha)\}$ and $\lim_{i\to\infty} \mathrm{Cr}\{\xi \geq x_i\} \geq \alpha > 0.5$, it follows from the credibility semicontinuity law that

$$\mathrm{Cr}\{\xi \geq \xi_{\sup}(\alpha)\} = \lim_{i\to\infty} \mathrm{Cr}\{\xi \geq x_i\} \geq \alpha.$$

The theorem is proved.

Theorem B.26 *Let ξ be a fuzzy variable, and $\alpha \in (0,1]$. Then we have*
(a) $\xi_{\inf}(\alpha)$ is an increasing and left-continuous function of α;
(b) $\xi_{\sup}(\alpha)$ is a decreasing and left-continuous function of α.

Proof: (a) Let α_1 and α_2 be two numbers with $0 < \alpha_1 < \alpha_2 \leq 1$. Then for any number $r < \xi_{\sup}(\alpha_2)$, we have

$$\mathrm{Cr}\left\{\xi \geq r\right\} \geq \alpha_2 > \alpha_1.$$

Thus, by the definition of optimistic value, we obtain $\xi_{\sup}(\alpha_1) \geq \xi_{\sup}(\alpha_2)$. That is, the value $\xi_{\sup}(\alpha)$ is a decreasing function of α. Next, we prove the left-continuity of $\xi_{\inf}(\alpha)$ with respect to α. Let $\{\alpha_i\}$ be an arbitrary sequence of positive numbers such that $\alpha_i \uparrow \alpha$. Then $\{\xi_{\inf}(\alpha_i)\}$ is an increasing sequence. If the limitation is equal to $\xi_{\inf}(\alpha)$, then the left-continuity is proved. Otherwise, there exists a number z^* such that

$$\lim_{i\to\infty} \xi_{\inf}(\alpha_i) < z^* < \xi_{\inf}(\alpha).$$

Thus $\mathrm{Cr}\{\xi \leq z^*\} \geq \alpha_i$ for each i. Letting $i \to \infty$, we get $\mathrm{Cr}\{\xi \leq z^*\} \geq \alpha$. Hence $z^* \geq \xi_{\inf}(\alpha)$. A contradiction proves the left-continuity of $\xi_{\inf}(\alpha)$ with respect to α. The part (b) may be proved similarly.

B.11 Entropy

Fuzzy entropy is a measure of uncertainty and has been studied by many researchers such as De Luca and Termini [27], Kaufmann [71], Yager [217], Kosko [80], Pal and Pal [173], Bhandari and Pal [7], and Pal and Bezdek [175]. Those definitions of entropy characterize the uncertainty resulting primarily from the linguistic vagueness rather than resulting from information deficiency, and vanishes when the fuzzy variable is an equipossible one.

Liu [126] suggested that an entropy of fuzzy variables should meet at least the following three basic requirements: (i) *minimum:* the entropy of a crisp number is minimum, i.e., 0; (ii) *maximum:* the entropy of an equipossible fuzzy variable is maximum; (iii) *universality:* the entropy is applicable not only to finite and infinite cases but also to discrete and continuous cases.

In order to meet those requirements, Li and Liu [91] provided a new definition of fuzzy entropy to characterize the uncertainty resulting from information deficiency which is caused by the impossibility to predict the specified value that a fuzzy variable takes.

Entropy of Discrete Fuzzy Variables

Definition B.15 *(Li and Liu [91]) Let ξ be a discrete fuzzy variable taking values in $\{x_1, x_2, \cdots\}$. Then its entropy is defined by*

$$H[\xi] = \sum_{i=1}^{\infty} S(\mathrm{Cr}\{\xi = x_i\}) \qquad (B.56)$$

where $S(t) = -t \ln t - (1-t) \ln(1-t)$.

Example B.30: Suppose that ξ is a discrete fuzzy variable taking values in $\{x_1, x_2, \cdots\}$. If there exists some index k such that the membership function $\mu(x_k) = 1$, and 0 otherwise, then its entropy $H[\xi] = 0$.

Example B.31: Suppose that ξ is a simple fuzzy variable taking values in $\{x_1, x_2, \cdots, x_n\}$. If its membership function $\mu(x) \equiv 1$, then its entropy $H[\xi] = n \ln 2$.

Theorem B.27 *Suppose that ξ is a discrete fuzzy variable taking values in $\{x_1, x_2, \cdots\}$. Then*
$$H[\xi] \geq 0 \qquad (B.57)$$

and equality holds if and only if ξ is essentially a crisp number.

Proof: The nonnegativity is clear. In addition, $H[\xi] = 0$ if and only if $\mathrm{Cr}\{\xi = x_i\} = 0$ or 1 for each i. That is, there exists one and only one index k such that $\mathrm{Cr}\{\xi = x_k\} = 1$, i.e., ξ is essentially a crisp number.

This theorem states that the entropy of a fuzzy variable reaches its minimum 0 when the fuzzy variable degenerates to a crisp number. In this case, there is no uncertainty.

Theorem B.28 *Suppose that ξ is a simple fuzzy variable taking values in $\{x_1, x_2, \cdots, x_n\}$. Then*
$$H[\xi] \leq n \ln 2 \qquad (B.58)$$

and equality holds if and only if ξ is an equipossible fuzzy variable.

Proof: Since the function $S(t)$ reaches its maximum $\ln 2$ at $t = 0.5$, we have

$$H[\xi] = \sum_{i=1}^{n} S(\mathrm{Cr}\{\xi = x_i\}) \leq n \ln 2$$

and equality holds if and only if $\mathrm{Cr}\{\xi = x_i\} = 0.5$, i.e., $\mu(x_i) \equiv 1$ for all $i = 1, 2, \cdots, n$.

This theorem states that the entropy of a fuzzy variable reaches its maximum when the fuzzy variable is an equipossible one. In this case, there is no preference among all the values that the fuzzy variable will take.

Entropy of Continuous Fuzzy Variables

Definition B.16 *(Li and Liu [91]) Let ξ be a continuous fuzzy variable. Then its entropy is defined by*

$$H[\xi] = \int_{-\infty}^{+\infty} S(\mathrm{Cr}\{\xi = x\})\mathrm{d}x \qquad (B.59)$$

where $S(t) = -t \ln t - (1-t) \ln(1-t)$.

For any continuous fuzzy variable ξ with membership function μ, we have $\mathrm{Cr}\{\xi = x\} = \mu(x)/2$ for each $x \in \Re$. Thus

$$H[\xi] = -\int_{-\infty}^{+\infty} \left(\frac{\mu(x)}{2} \ln \frac{\mu(x)}{2} + \left(1 - \frac{\mu(x)}{2}\right) \ln \left(1 - \frac{\mu(x)}{2}\right) \right) \mathrm{d}x. \quad (B.60)$$

Example B.32: Let ξ be an equipossible fuzzy variable (a,b). Then $\mu(x) = 1$ if $a \le x \le b$, and 0 otherwise. Thus its entropy is

$$H[\xi] = -\int_{a}^{b} \left(\frac{1}{2} \ln \frac{1}{2} + \left(1 - \frac{1}{2}\right) \ln \left(1 - \frac{1}{2}\right) \right) \mathrm{d}x = (b-a)\ln 2.$$

Example B.33: Let ξ be a triangular fuzzy variable (a,b,c). Then its entropy is $H[\xi] = (c-a)/2$.

Example B.34: Let ξ be a trapezoidal fuzzy variable (a,b,c,d). Then its entropy is $H[\xi] = (d-a)/2 + (\ln 2 - 0.5)(c-b)$.

Example B.35: Let ξ be an exponentially distributed fuzzy variable with second moment m^2. Then its entropy is $H[\xi] = \pi m/\sqrt{6}$.

Example B.36: Let ξ be a normally distributed fuzzy variable with expected value e and variance σ^2. Then its entropy is $H[\xi] = \sqrt{6}\pi\sigma/3$.

Theorem B.29 *Let ξ be a continuous fuzzy variable taking values on the interval $[a,b]$. Then*

$$H[\xi] \le (b-a)\ln 2 \qquad (B.61)$$

and equality holds if and only if ξ is an equipossible fuzzy variable (a,b).

Proof: The theorem follows from the fact that the function $S(t)$ reaches its maximum $\ln 2$ at $t = 0.5$.

Maximum Entropy Principle

Given some constraints, for example, expected value and variance, there are usually multiple compatible membership functions. Which membership function shall we take? The *maximum entropy principle* attempts to select the membership function that maximizes the value of entropy and satisfies the prescribed constraints.

Theorem B.30 *(Li and Liu [95]) Let ξ be a continuous nonnegative fuzzy variable with finite second moment m^2. Then*

$$H[\xi] \leq \frac{\pi m}{\sqrt{6}} \tag{B.62}$$

and the equality holds if ξ is an exponentially distributed fuzzy variable with second moment m^2.

Proof: Let μ be the membership function of ξ. Note that μ is a continuous function. The proof is based on the following two steps.

Step 1: Suppose that μ is a decreasing function on $[0, +\infty)$. For this case, we have $\text{Cr}\{\xi \geq x\} = \mu(x)/2$ for any $x > 0$. Thus the second moment

$$E[\xi^2] = \int_0^{+\infty} \text{Cr}\{\xi^2 \geq x\}\mathrm{d}x = \int_0^{+\infty} 2x\text{Cr}\{\xi \geq x\}\mathrm{d}x = \int_0^{+\infty} x\mu(x)\mathrm{d}x.$$

The maximum entropy membership function μ should maximize the entropy

$$-\int_0^{+\infty} \left(\frac{\mu(x)}{2} \ln \frac{\mu(x)}{2} + \left(1 - \frac{\mu(x)}{2}\right) \ln \left(1 - \frac{\mu(x)}{2}\right) \right) \mathrm{d}x$$

subject to the moment constraint

$$\int_0^{+\infty} x\mu(x)\mathrm{d}x = m^2.$$

The Lagrangian is

$$L = -\int_0^{+\infty} \left(\frac{\mu(x)}{2} \ln \frac{\mu(x)}{2} + \left(1 - \frac{\mu(x)}{2}\right) \ln \left(1 - \frac{\mu(x)}{2}\right) \right) \mathrm{d}x$$

$$- \lambda \left(\int_0^{+\infty} x\mu(x)\mathrm{d}x - m^2 \right).$$

The maximum entropy membership function meets Euler-Lagrange equation

$$\frac{1}{2} \ln \frac{\mu(x)}{2} - \frac{1}{2} \ln \left(1 - \frac{\mu(x)}{2}\right) + \lambda x = 0$$

and has the form $\mu(x) = 2\left(1 + \exp(2\lambda x)\right)^{-1}$. Substituting it into the moment constraint, we get

$$\mu^*(x) = 2\left(1 + \exp\left(\frac{\pi x}{\sqrt{6}m}\right)\right)^{-1}, \quad x \geq 0$$

which is just the exponential membership function with second moment m^2, and the maximum entropy is $H[\xi^*] = \pi m/\sqrt{6}$.

STEP 2: Let ξ be a general fuzzy variable with second moment m^2. Now we define a fuzzy variable $\widehat{\xi}$ via membership function

$$\widehat{\mu}(x) = \sup_{y \geq x} \mu(y), \quad x \geq 0.$$

Then $\widehat{\mu}$ is a decreasing function on $[0, +\infty)$, and

$$\mathrm{Cr}\{\widehat{\xi}^2 \geq x\} = \frac{1}{2} \sup_{y \geq \sqrt{x}} \widehat{\mu}(y) = \frac{1}{2} \sup_{y \geq \sqrt{x}} \sup_{z \geq y} \mu(z) = \frac{1}{2} \sup_{z \geq \sqrt{x}} \mu(z) \leq \mathrm{Cr}\{\xi^2 \geq x\}$$

for any $x > 0$. Thus we have

$$E[\widehat{\xi}^2] = \int_0^{+\infty} \mathrm{Cr}\{\widehat{\xi}^2 \geq x\}\mathrm{d}x \leq \int_0^{+\infty} \mathrm{Cr}\{\xi^2 \geq x\}\mathrm{d}x = E[\xi^2] = m^2.$$

It follows from $\mu(x) \leq \widehat{\mu}(x)$ and Step 1 that

$$H[\xi] \leq H[\widehat{\xi}] \leq \frac{\pi \sqrt{E[\widehat{\xi}^2]}}{\sqrt{6}} \leq \frac{\pi m}{\sqrt{6}}.$$

The theorem is thus proved.

Theorem B.31 *(Li and Liu [95]) Let ξ be a continuous fuzzy variable with finite expected value e and variance σ^2. Then*

$$H[\xi] \leq \frac{\sqrt{6}\pi\sigma}{3} \tag{B.63}$$

and the equality holds if ξ is a normally distributed fuzzy variable with expected value e and variance σ^2.

Proof: Let μ be the continuous membership function of ξ. The proof is based on the following two steps.

STEP 1: Let $\mu(x)$ be a unimodal and symmetric function about $x = e$. For this case, the variance is

$$V[\xi] = \int_0^{+\infty} \mathrm{Cr}\{(\xi - e)^2 \geq x\}\mathrm{d}x = \int_0^{+\infty} \mathrm{Cr}\{\xi - e \geq \sqrt{x}\}\mathrm{d}x$$

$$= \int_e^{+\infty} 2(x - e)\mathrm{Cr}\{\xi \geq x\}\mathrm{d}x = \int_e^{+\infty} (x - e)\mu(x)\mathrm{d}x$$

and the entropy is

$$H[\xi] = -2 \int_e^{+\infty} \left(\frac{\mu(x)}{2} \ln \frac{\mu(x)}{2} + \left(1 - \frac{\mu(x)}{2}\right) \ln \left(1 - \frac{\mu(x)}{2}\right) \right) \mathrm{d}x.$$

The maximum entropy membership function μ should maximize the entropy subject to the variance constraint. The Lagrangian is

$$L = -2 \int_e^{+\infty} \left(\frac{\mu(x)}{2} \ln \frac{\mu(x)}{2} + \left(1 - \frac{\mu(x)}{2}\right) \ln \left(1 - \frac{\mu(x)}{2}\right) \right) \mathrm{d}x$$

$$-\lambda \left(\int_e^{+\infty} (x - e)\mu(x)\mathrm{d}x - \sigma^2 \right).$$

The maximum entropy membership function meets Euler-Lagrange equation

$$\ln \frac{\mu(x)}{2} - \ln \left(1 - \frac{\mu(x)}{2}\right) + \lambda(x - e) = 0$$

and has the form $\mu(x) = 2 \left(1 + \exp \left(\lambda(x - e)\right)\right)^{-1}$. Substituting it into the variance constraint, we get

$$\mu^*(x) = 2 \left(1 + \exp \left(\frac{\pi|x - e|}{\sqrt{6}\sigma}\right)\right)^{-1}, \quad x \in \Re$$

which is just the normal membership function with expected value e and variance σ^2, and the maximum entropy is $H[\xi^*] = \sqrt{6}\pi\sigma/3$.

STEP 2: Let ξ be a general fuzzy variable with expected value e and variance σ^2. We define a fuzzy variable $\widehat{\xi}$ by the membership function

$$\widehat{\mu}(x) = \begin{cases} \sup_{y \leq x}(\mu(y) \vee \mu(2e - y)), & \text{if } x \leq e \\ \sup_{y \geq x} (\mu(y) \vee \mu(2e - y)), & \text{if } x > e. \end{cases}$$

It is easy to verify that $\widehat{\mu}(x)$ is a unimodal and symmetric function about $x = e$. Furthermore,

$$\mathrm{Cr} \left\{ (\widehat{\xi} - e)^2 \geq r \right\} = \frac{1}{2} \sup_{x \geq e+\sqrt{r}} \widehat{\mu}(x) = \frac{1}{2} \sup_{x \geq e+\sqrt{r}} \sup_{y \geq x}(\mu(y) \vee \mu(2e - y))$$

$$= \frac{1}{2} \sup_{y \geq e+\sqrt{r}} (\mu(y) \vee \mu(2e - y)) = \frac{1}{2} \sup_{(y-e)^2 \geq r} \mu(y)$$

$$\leq \mathrm{Cr} \left\{ (\xi - e)^2 \geq r \right\}$$

for any $r > 0$. Thus

$$V[\widehat{\xi}] = \int_0^{+\infty} \mathrm{Cr}\{(\widehat{\xi} - e)^2 \geq r\}\mathrm{d}r \leq \int_0^{+\infty} \mathrm{Cr}\{(\xi - e)^2 \geq r\}\mathrm{d}r = \sigma^2.$$

It follows from $\mu(x) \leq \widehat{\mu}(x)$ and Step 1 that

$$H[\xi] \leq H[\widehat{\xi}] \leq \frac{\sqrt{6}\pi\sqrt{V[\widehat{\xi}]}}{3} \leq \frac{\sqrt{6}\pi\sigma}{3}.$$

The proof is complete.

B.12   Conditional Credibility

We now consider the credibility of an event A after it has been learned that some other event B has occurred. This new credibility of A is called the conditional credibility of A given B.

In order to define a conditional credibility measure $\text{Cr}\{A|B\}$, at first we have to enlarge $\text{Cr}\{A \cap B\}$ because $\text{Cr}\{A \cap B\} < 1$ for all events whenever $\text{Cr}\{B\} < 1$. It seems that we have no alternative but to divide $\text{Cr}\{A \cap B\}$ by $\text{Cr}\{B\}$. Unfortunately, $\text{Cr}\{A \cap B\}/\text{Cr}\{B\}$ is not always a credibility measure. However, the value $\text{Cr}\{A|B\}$ should not be greater than $\text{Cr}\{A \cap B\}/\text{Cr}\{B\}$ (otherwise the normality will be lost), i.e.,

$$\text{Cr}\{A|B\} \leq \frac{\text{Cr}\{A \cap B\}}{\text{Cr}\{B\}}. \tag{B.64}$$

On the other hand, in order to preserve the self-duality, we should have

$$\text{Cr}\{A|B\} = 1 - \text{Cr}\{A^c|B\} \geq 1 - \frac{\text{Cr}\{A^c \cap B\}}{\text{Cr}\{B\}}. \tag{B.65}$$

Furthermore, since $(A \cap B) \cup (A^c \cap B) = B$, we have $\text{Cr}\{B\} \leq \text{Cr}\{A \cap B\} + \text{Cr}\{A^c \cap B\}$ by using the credibility subadditivity theorem. Thus

$$0 \leq 1 - \frac{\text{Cr}\{A^c \cap B\}}{\text{Cr}\{B\}} \leq \frac{\text{Cr}\{A \cap B\}}{\text{Cr}\{B\}} \leq 1. \tag{B.66}$$

Hence any numbers between $1 - \text{Cr}\{A^c \cap B\}/\text{Cr}\{B\}$ and $\text{Cr}\{A \cap B\}/\text{Cr}\{B\}$ are reasonable values that the conditional credibility may take. Based on the maximum uncertainty principle, we have the following conditional credibility measure.

Definition B.17 *(Liu [127]) Let $(\Theta, \mathcal{P}, \text{Cr})$ be a credibility space, and $A, B \in \mathcal{P}$. Then the conditional credibility measure of A given B is defined by*

$$\text{Cr}\{A|B\} = \begin{cases} \dfrac{\text{Cr}\{A \cap B\}}{\text{Cr}\{B\}}, & \text{if } \dfrac{\text{Cr}\{A \cap B\}}{\text{Cr}\{B\}} < 0.5 \\[2ex] 1 - \dfrac{\text{Cr}\{A^c \cap B\}}{\text{Cr}\{B\}}, & \text{if } \dfrac{\text{Cr}\{A^c \cap B\}}{\text{Cr}\{B\}} < 0.5 \\[2ex] 0.5, & \text{otherwise} \end{cases} \tag{B.67}$$

provided that $\text{Cr}\{B\} > 0$.

It follows immediately from the definition of conditional credibility that

$$1 - \frac{\text{Cr}\{A^c \cap B\}}{\text{Cr}\{B\}} \leq \text{Cr}\{A|B\} \leq \frac{\text{Cr}\{A \cap B\}}{\text{Cr}\{B\}}. \tag{B.68}$$

Furthermore, the value of $\text{Cr}\{A|B\}$ takes values as close to 0.5 as possible in the interval. In other words, it accords with the maximum uncertainty principle.

Theorem B.32 *Let $(\Theta, \mathcal{P}, \mathrm{Cr})$ be a credibility space, and B an event with $\mathrm{Cr}\{B\} > 0$. Then $\mathrm{Cr}\{\cdot|B\}$ defined by (B.67) is a credibility measure, and $(\Theta, \mathcal{P}, \mathrm{Cr}\{\cdot|B\})$ is a credibility space.*

Proof: It is sufficient to prove that $\mathrm{Cr}\{\cdot|B\}$ satisfies the normality, monotonicity, self-duality and maximality axioms. At first, it satisfies the normality axiom, i.e.,

$$\mathrm{Cr}\{\Theta|B\} = 1 - \frac{\mathrm{Cr}\{\Theta^c \cap B\}}{\mathrm{Cr}\{B\}} = 1 - \frac{\mathrm{Cr}\{\emptyset\}}{\mathrm{Cr}\{B\}} = 1.$$

For any events A_1 and A_2 with $A_1 \subset A_2$, if

$$\frac{\mathrm{Cr}\{A_1 \cap B\}}{\mathrm{Cr}\{B\}} \leq \frac{\mathrm{Cr}\{A_2 \cap B\}}{\mathrm{Cr}\{B\}} < 0.5,$$

then

$$\mathrm{Cr}\{A_1|B\} = \frac{\mathrm{Cr}\{A_1 \cap B\}}{\mathrm{Cr}\{B\}} \leq \frac{\mathrm{Cr}\{A_2 \cap B\}}{\mathrm{Cr}\{B\}} = \mathrm{Cr}\{A_2|B\}.$$

If

$$\frac{\mathrm{Cr}\{A_1 \cap B\}}{\mathrm{Cr}\{B\}} \leq 0.5 \leq \frac{\mathrm{Cr}\{A_2 \cap B\}}{\mathrm{Cr}\{B\}},$$

then $\mathrm{Cr}\{A_1|B\} \leq 0.5 \leq \mathrm{Cr}\{A_2|B\}$. If

$$0.5 < \frac{\mathrm{Cr}\{A_1 \cap B\}}{\mathrm{Cr}\{B\}} \leq \frac{\mathrm{Cr}\{A_2 \cap B\}}{\mathrm{Cr}\{B\}},$$

then we have

$$\mathrm{Cr}\{A_1|B\} = \left(1 - \frac{\mathrm{Cr}\{A_1^c \cap B\}}{\mathrm{Cr}\{B\}}\right) \vee 0.5 \leq \left(1 - \frac{\mathrm{Cr}\{A_2^c \cap B\}}{\mathrm{Cr}\{B\}}\right) \vee 0.5 = \mathrm{Cr}\{A_2|B\}.$$

This means that $\mathrm{Cr}\{\cdot|B\}$ satisfies the monotonicity axiom. For any event A, if

$$\frac{\mathrm{Cr}\{A \cap B\}}{\mathrm{Cr}\{B\}} \geq 0.5, \quad \frac{\mathrm{Cr}\{A^c \cap B\}}{\mathrm{Cr}\{B\}} \geq 0.5,$$

then we have $\mathrm{Cr}\{A|B\} + \mathrm{Cr}\{A^c|B\} = 0.5 + 0.5 = 1$ immediately. Otherwise, without loss of generality, suppose

$$\frac{\mathrm{Cr}\{A \cap B\}}{\mathrm{Cr}\{B\}} < 0.5 < \frac{\mathrm{Cr}\{A^c \cap B\}}{\mathrm{Cr}\{B\}},$$

then we have

$$\mathrm{Cr}\{A|B\} + \mathrm{Cr}\{A^c|B\} = \frac{\mathrm{Cr}\{A \cap B\}}{\mathrm{Cr}\{B\}} + \left(1 - \frac{\mathrm{Cr}\{A \cap B\}}{\mathrm{Cr}\{B\}}\right) = 1.$$

That is, $\mathrm{Cr}\{\cdot|B\}$ satisfies the self-duality axiom. Finally, for any events $\{A_i\}$ with $\sup_i \mathrm{Cr}\{A_i|B\} < 0.5$, we have $\sup_i \mathrm{Cr}\{A_i \cap B\} < 0.5$ and

$$\sup_i \mathrm{Cr}\{A_i|B\} = \frac{\sup_i \mathrm{Cr}\{A_i \cap B\}}{\mathrm{Cr}\{B\}} = \frac{\mathrm{Cr}\{\cup_i A_i \cap B\}}{\mathrm{Cr}\{B\}} = \mathrm{Cr}\{\cup_i A_i|B\}.$$

Thus $\mathrm{Cr}\{\cdot|B\}$ satisfies the maximality axiom. Hence $\mathrm{Cr}\{\cdot|B\}$ is a credibility measure. Furthermore, $(\Theta, \mathcal{P}, \mathrm{Cr}\{\cdot|B\})$ is a credibility space.

Definition B.18 *(Liu [127]) The conditional membership function of a fuzzy variable ξ given B is defined by*

$$\mu(x|B) = (2\mathrm{Cr}\{\xi = x|B\}) \wedge 1, \quad x \in \Re \tag{B.69}$$

provided that $\mathrm{Cr}\{B\} > 0$.

Example B.37: Let ξ be a fuzzy variable with membership function $\mu(x)$, and X a set of real numbers such that $\mu(x) > 0$ for some $x \in X$. Then the conditional membership function of ξ given $\xi \in X$ is

$$\mu(x|X) = \begin{cases} \dfrac{2\mu(x)}{\sup\limits_{x \in X} \mu(x)} \wedge 1, & \text{if } \sup\limits_{x \in X} \mu(x) < 1 \\[3ex] \dfrac{2\mu(x)}{2 - \sup\limits_{x \in X^c} \mu(x)} \wedge 1, & \text{if } \sup\limits_{x \in X} \mu(x) = 1 \end{cases} \tag{B.70}$$

for $x \in X$. Please mention that $\mu(x|X) \equiv 0$ if $x \notin X$.
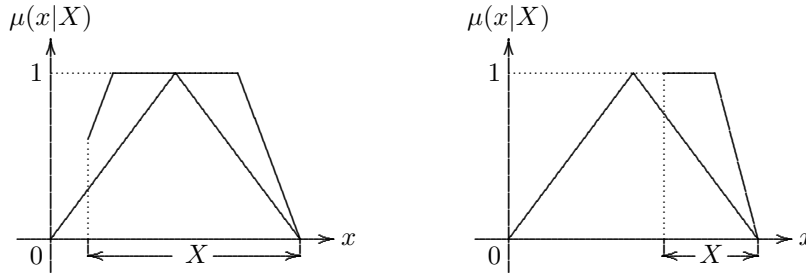
Figure B.5: Conditional Membership Function $\mu(x|X)$

Example B.38: Let ξ and η be two fuzzy variables with joint membership function $\mu(x, y)$, and Y a set of real numbers. Then the conditional

membership function of ξ given $\eta \in Y$ is

$$
\mu(x|Y) =
\begin{cases}
\dfrac{2 \sup\limits_{y \in Y} \mu(x,y)}{\sup\limits_{x \in \Re, y \in Y} \mu(x,y)} \wedge 1, & \text{if } \sup\limits_{x \in \Re, y \in Y} \mu(x,y) < 1 \\[4ex]
\dfrac{2 \sup\limits_{y \in Y} \mu(x,y)}{2 - \sup\limits_{x \in \Re, y \in Y^c} \mu(x,y)} \wedge 1, & \text{if } \sup\limits_{x \in \Re, y \in Y} \mu(x,y) = 1
\end{cases}
\tag{B.71}
$$

provided that $\mu(x,y) > 0$ for some $x \in \Re$ and $y \in Y$. Especially, the conditional membership function of ξ given $\eta = y$ is

$$
\mu(x|y) =
\begin{cases}
\dfrac{2\mu(x,y)}{\sup\limits_{x \in \Re} \mu(x,y)} \wedge 1, & \text{if } \sup\limits_{x \in \Re} \mu(x,y) < 1 \\[4ex]
\dfrac{2\mu(x,y)}{2 - \sup\limits_{x \in \Re, z \neq y} \mu(x,z)} \wedge 1, & \text{if } \sup\limits_{x \in \Re} \mu(x,y) = 1
\end{cases}
$$

provided that $\mu(x,y) > 0$ for some $x \in \Re$.

Definition B.19 *Let ξ be a fuzzy variable on $(\Theta, \mathcal{P}, \mathrm{Cr})$. A conditional fuzzy variable of ξ given B is a function $\xi|_B$ from the conditional credibility space $(\Theta, \mathcal{P}, \mathrm{Cr}\{\cdot|B\})$ to the set of real numbers such that*

$$
\xi|_B(\theta) \equiv \xi(\theta), \quad \forall \theta \in \Theta.
\tag{B.72}
$$

Definition B.20 *(Liu [127]) The conditional credibility distribution $\Phi \colon \Re \to [0,1]$ of a fuzzy variable ξ given B is defined by*

$$
\Phi(x|B) = \mathrm{Cr}\left\{\xi \leq x|B\right\}
\tag{B.73}
$$

provided that $\mathrm{Cr}\{B\} > 0$.

Definition B.21 *(Liu [127]) Let ξ be a fuzzy variable. Then the conditional expected value of ξ given B is defined by*

$$
E[\xi|B] = \int_0^{+\infty} \mathrm{Cr}\{\xi \geq r|B\}\mathrm{d}r - \int_{-\infty}^0 \mathrm{Cr}\{\xi \leq r|B\}\mathrm{d}r
\tag{B.74}
$$

provided that at least one of the two integrals is finite.

Appendix C

Chance Theory

Fuzziness and randomness are two basic types of uncertainty. In many cases, fuzziness and randomness simultaneously appear in a system. In order to describe this phenomena, a fuzzy random variable was introduced by Kwakernaak [82] as a random element taking "fuzzy variable" values. In addition, a random fuzzy variable was proposed by Liu [119] as a fuzzy element taking "random variable" values. For example, it might be known that the lifetime of a modern engine is an exponentially distributed random variable with an unknown parameter. If the parameter is provided as a fuzzy variable, then the lifetime is a random fuzzy variable.

More generally, a hybrid variable was introduced by Liu [125] in 2006 as a tool to describe the quantities with fuzziness and randomness. Fuzzy random variable and random fuzzy variable are instances of hybrid variable. In order to measure hybrid events, a concept of chance measure was introduced by Li and Liu [98] in 2009. Chance theory is a hybrid of probability theory and credibility theory. Perhaps the reader would like to know what axioms we should assume for chance theory. In fact, chance theory will be based on the three axioms of probability and five axioms of credibility.

The emphasis in this appendix is mainly on chance space, hybrid variable, chance measure, chance distribution, expected value, variance, critical values and conditional chance.

C.1   Chance Space

Chance theory begins with the concept of chance space that inherits the mathematical foundations of both probability theory and credibility theory.

Definition C.1 *(Liu [125]) Suppose that $(\Theta, \mathcal{P}, \mathrm{Cr})$ is a credibility space and $(\Omega, \mathcal{A}, \mathrm{Pr})$ is a probability space. The product $(\Theta, \mathcal{P}, \mathrm{Cr}) \times (\Omega, \mathcal{A}, \mathrm{Pr})$ is called a chance space.*

The universal set $\Theta \times \Omega$ is clearly the set of all ordered pairs of the form (θ, ω), where $\theta \in \Theta$ and $\omega \in \Omega$. What is the product σ-algebra $\mathcal{P} \times \mathcal{A}$? What is the product measure $\mathrm{Cr} \times \mathrm{Pr}$? Let us discuss these two basic problems.

What is the product σ-algebra $\mathcal{P} \times \mathcal{A}$?

Generally speaking, it is not true that all subsets of $\Theta \times \Omega$ are measurable. Let Λ be a subset of $\Theta \times \Omega$. Write

$$\Lambda(\theta) = \left\{ \omega \in \Omega \,\big|\, (\theta, \omega) \in \Lambda \right\}. \tag{C.1}$$

It is clear that $\Lambda(\theta)$ is a subset of Ω. If $\Lambda(\theta) \in \mathcal{A}$ holds for each $\theta \in \Theta$, then Λ may be regarded as a *measurable set*.
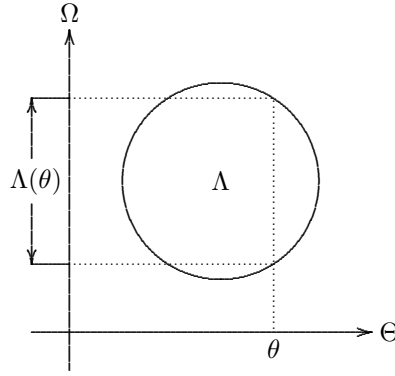
Figure C.1: $\Theta \times \Omega$, Λ and $\Lambda(\theta)$

Definition C.2 *(Liu [127]) Let $(\Theta, \mathcal{P}, \mathrm{Cr}) \times (\Omega, \mathcal{A}, \mathrm{Pr})$ be a chance space. A subset $\Lambda \subset \Theta \times \Omega$ is called an event if $\Lambda(\theta) \in \mathcal{A}$ for each $\theta \in \Theta$.*

Example C.1: Empty set $\emptyset$ and universal set $\Theta \times \Omega$ are clearly events.

Example C.2: Let $X \in \mathcal{P}$ and $Y \in \mathcal{A}$. Then $X \times Y$ is a subset of $\Theta \times \Omega$. Since the set

$$(X \times Y)(\theta) = \begin{cases} Y, & \text{if } \theta \in X \\ \emptyset, & \text{if } \theta \in X^c \end{cases}$$

is in the σ-algebra $\mathcal{A}$ for each $\theta \in \Theta$, the rectangle $X \times Y$ is an event.

Theorem C.1 *(Liu [127]) Let $(\Theta, \mathcal{P}, \mathrm{Cr}) \times (\Omega, \mathcal{A}, \mathrm{Pr})$ be a chance space. The class of all events is a σ-algebra over $\Theta \times \Omega$, and denoted by $\mathcal{P} \times \mathcal{A}$.*

Proof: At first, it is obvious that $\Theta \times \Omega \in \mathcal{P} \times \mathcal{A}$. For any event Λ, we always have

$$\Lambda(\theta) \in \mathcal{A}, \quad \forall \theta \in \Theta.$$

Thus for each $\theta \in \Theta$, the set

$$\Lambda^c(\theta) = \left\{ \omega \in \Omega \mid (\theta, \omega) \in \Lambda^c \right\} = (\Lambda(\theta))^c \in \mathcal{A}$$

which implies that $\Lambda^c \in \mathcal{P} \times \mathcal{A}$. Finally, let $\Lambda_1, \Lambda_2, \cdots$ be events. Then for each $\theta \in \Theta$, we have

$$\left(\bigcup_{i=1}^{\infty} \Lambda_i \right)(\theta) = \left\{ \omega \in \Omega \mid (\theta, \omega) \in \bigcup_{i=1}^{\infty} \Lambda_i \right\} = \bigcup_{i=1}^{\infty} \left\{ \omega \in \Omega \mid (\theta, \omega) \in \Lambda_i \right\} \in \mathcal{A}.$$

That is, the countable union $\cup_i \Lambda_i \in \mathcal{P} \times \mathcal{A}$. Hence $\mathcal{P} \times \mathcal{A}$ is a σ-algebra.

What is the product measure $\mathrm{Cr} \times \mathrm{Pr}$**?**

Product probability is a probability measure, and product credibility is a credibility measure. What is the product measure $\mathrm{Cr} \times \mathrm{Pr}$? We will call it *chance measure* and define it as follows.

Definition C.3 *(Li and Liu [98]) Let* $(\Theta, \mathcal{P}, \mathrm{Cr}) \times (\Omega, \mathcal{A}, \mathrm{Pr})$ *be a chance space. Then a chance measure of an event* Λ *is defined as*

$$\mathrm{Ch}\{\Lambda\} = \begin{cases} \sup_{\theta \in \Theta}(\mathrm{Cr}\{\theta\} \wedge \mathrm{Pr}\{\Lambda(\theta)\}), \\ \qquad \textit{if } \sup_{\theta \in \Theta}(\mathrm{Cr}\{\theta\} \wedge \mathrm{Pr}\{\Lambda(\theta)\}) < 0.5 \\ 1 - \sup_{\theta \in \Theta}(\mathrm{Cr}\{\theta\} \wedge \mathrm{Pr}\{\Lambda^c(\theta)\}), \\ \qquad \textit{if } \sup_{\theta \in \Theta}(\mathrm{Cr}\{\theta\} \wedge \mathrm{Pr}\{\Lambda(\theta)\}) \geq 0.5. \end{cases} \quad (C.2)$$

Example C.3: Take a credibility space $(\Theta, \mathcal{P}, \mathrm{Cr})$ to be $\{\theta_1, \theta_2\}$ with $\mathrm{Cr}\{\theta_1\} = 0.6$ and $\mathrm{Cr}\{\theta_2\} = 0.4$, and take a probability space $(\Omega, \mathcal{A}, \mathrm{Pr})$ to be $\{\omega_1, \omega_2\}$ with $\mathrm{Pr}\{\omega_1\} = 0.7$ and $\mathrm{Pr}\{\omega_2\} = 0.3$. Then

$$\mathrm{Ch}\{(\theta_1, \omega_1)\} = 0.6, \quad \mathrm{Ch}\{(\theta_2, \omega_2)\} = 0.3.$$

Theorem C.2 *Let* $(\Theta, \mathcal{P}, \mathrm{Cr}) \times (\Omega, \mathcal{A}, \mathrm{Pr})$ *be a chance space and* Ch *a chance measure. Then we have*

$$\mathrm{Ch}\{\emptyset\} = 0, \quad (C.3)$$

$$\mathrm{Ch}\{\Theta \times \Omega\} = 1, \quad (C.4)$$

$$0 \leq \mathrm{Ch}\{\Lambda\} \leq 1 \quad (C.5)$$

for any event Λ.

Proof: It follows from the definition immediately.

Theorem C.3 *Let* $(\Theta, \mathcal{P}, \mathrm{Cr}) \times (\Omega, \mathcal{A}, \mathrm{Pr})$ *be a chance space and* Ch *a chance measure. Then for any event* Λ, *we have*

$$\sup_{\theta \in \Theta}(\mathrm{Cr}\{\theta\} \wedge \mathrm{Pr}\{\Lambda(\theta)\}) \vee \sup_{\theta \in \Theta}(\mathrm{Cr}\{\theta\} \wedge \mathrm{Pr}\{\Lambda^c(\theta)\}) \geq 0.5, \qquad (C.6)$$

$$\sup_{\theta \in \Theta}(\mathrm{Cr}\{\theta\} \wedge \mathrm{Pr}\{\Lambda(\theta)\}) + \sup_{\theta \in \Theta}(\mathrm{Cr}\{\theta\} \wedge \mathrm{Pr}\{\Lambda^c(\theta)\}) \leq 1, \qquad (C.7)$$

$$\sup_{\theta \in \Theta}(\mathrm{Cr}\{\theta\} \wedge \mathrm{Pr}\{\Lambda(\theta)\}) \leq \mathrm{Ch}\{\Lambda\} \leq 1 - \sup_{\theta \in \Theta}(\mathrm{Cr}\{\theta\} \wedge \mathrm{Pr}\{\Lambda^c(\theta)\}). \quad (C.8)$$

Proof: It follows from the basic properties of probability and credibility that

$$\sup_{\theta \in \Theta}(\mathrm{Cr}\{\theta\} \wedge \mathrm{Pr}\{\Lambda(\theta)\}) \vee \sup_{\theta \in \Theta}(\mathrm{Cr}\{\theta\} \wedge \mathrm{Pr}\{\Lambda^c(\theta)\})$$

$$\geq \sup_{\theta \in \Theta}(\mathrm{Cr}\{\theta\} \wedge (\mathrm{Pr}\{\Lambda(\theta)\} \vee \mathrm{Pr}\{\Lambda^c(\theta)\}))$$

$$\geq \sup_{\theta \in \Theta} \mathrm{Cr}\{\theta\} \wedge 0.5 = 0.5$$

and

$$\sup_{\theta \in \Theta}(\mathrm{Cr}\{\theta\} \wedge \mathrm{Pr}\{\Lambda(\theta)\}) + \sup_{\theta \in \Theta}(\mathrm{Cr}\{\theta\} \wedge \mathrm{Pr}\{\Lambda^c(\theta)\})$$

$$= \sup_{\theta_1, \theta_2 \in \Theta} (\mathrm{Cr}\{\theta_1\} \wedge \mathrm{Pr}\{\Lambda(\theta_1)\} + \mathrm{Cr}\{\theta_2\} \wedge \mathrm{Pr}\{\Lambda^c(\theta_2)\})$$

$$\leq \sup_{\theta_1 \neq \theta_2} (\mathrm{Cr}\{\theta_1\} + \mathrm{Cr}\{\theta_2\}) \vee \sup_{\theta \in \Theta}(\mathrm{Pr}\{\Lambda(\theta)\} + \mathrm{Pr}\{\Lambda^c(\theta)\})$$

$$\leq 1 \vee 1 = 1.$$

The inequalities (C.8) follows immediately from the above inequalities and the definition of chance measure.

Theorem C.4 *(Li and Liu [98]) The chance measure is increasing. That is,*

$$\mathrm{Ch}\{\Lambda_1\} \leq \mathrm{Ch}\{\Lambda_2\} \qquad (C.9)$$

for any events Λ_1 *and* Λ_2 *with* $\Lambda_1 \subset \Lambda_2$.

Proof: Since $\Lambda_1(\theta) \subset \Lambda_2(\theta)$ and $\Lambda_2^c(\theta) \subset \Lambda_1^c(\theta)$ for each $\theta \in \Theta$, we have

$$\sup_{\theta \in \Theta}(\mathrm{Cr}\{\theta\} \wedge \mathrm{Pr}\{\Lambda_1(\theta)\}) \leq \sup_{\theta \in \Theta}(\mathrm{Cr}\{\theta\} \wedge \mathrm{Pr}\{\Lambda_2(\theta)\}),$$

$$\sup_{\theta \in \Theta}(\mathrm{Cr}\{\theta\} \wedge \mathrm{Pr}\{\Lambda_2^c(\theta)\}) \leq \sup_{\theta \in \Theta}(\mathrm{Cr}\{\theta\} \wedge \mathrm{Pr}\{\Lambda_1^c(\theta)\}).$$

The argument breaks down into three cases.

Case 1: $\sup_{\theta \in \Theta}(\mathrm{Cr}\{\theta\} \wedge \mathrm{Pr}\{\Lambda_2(\theta)\}) < 0.5$. For this case, we have

$$\sup_{\theta \in \Theta}(\mathrm{Cr}\{\theta\} \wedge \mathrm{Pr}\{\Lambda_1(\theta)\}) < 0.5,$$

$$\mathrm{Ch}\{\Lambda_2\} = \sup_{\theta\in\Theta}(\mathrm{Cr}\{\theta\}\wedge\mathrm{Pr}\{\Lambda_2(\theta)\}) \geq \sup_{\theta\in\Theta}(\mathrm{Cr}\{\theta\}\wedge\mathrm{Pr}\{\Lambda_1(\theta)\} = \mathrm{Ch}\{\Lambda_1\}.$$

Case 2: $\sup\limits_{\theta\in\Theta}(\mathrm{Cr}\{\theta\}\wedge\mathrm{Pr}\{\Lambda_2(\theta)\}) \geq 0.5$ and $\sup\limits_{\theta\in\Theta}(\mathrm{Cr}\{\theta\}\wedge\mathrm{Pr}\{\Lambda_1(\theta)\}) < 0.5$. It follows from Theorem C.3 that

$$\mathrm{Ch}\{\Lambda_2\} \geq \sup_{\theta\in\Theta}(\mathrm{Cr}\{\theta\}\wedge\mathrm{Pr}\{\Lambda_2(\theta)\}) \geq 0.5 > \mathrm{Ch}\{\Lambda_1\}.$$

Case 3: $\sup\limits_{\theta\in\Theta}(\mathrm{Cr}\{\theta\}\wedge\mathrm{Pr}\{\Lambda_2(\theta)\}) \geq 0.5$ and $\sup\limits_{\theta\in\Theta}(\mathrm{Cr}\{\theta\}\wedge\mathrm{Pr}\{\Lambda_1(\theta)\}) \geq 0.5$. For this case, we have

$$\mathrm{Ch}\{\Lambda_2\} = 1-\sup_{\theta\in\Theta}(\mathrm{Cr}\{\theta\}\wedge\mathrm{Pr}\{\Lambda_2^c(\theta)\}) \geq 1-\sup_{\theta\in\Theta}(\mathrm{Cr}\{\theta\}\wedge\mathrm{Pr}\{\Lambda_1^c(\theta)\}) = \mathrm{Ch}\{\Lambda_1\}.$$

Thus Ch is an increasing measure.

Theorem C.5 *(Li and Liu [98]) The chance measure is self-dual. That is,*

$$\mathrm{Ch}\{\Lambda\} + \mathrm{Ch}\{\Lambda^c\} = 1 \qquad\qquad (C.10)$$

for any event Λ.

Proof: For any event Λ, please note that

$$\mathrm{Ch}\{\Lambda^c\} = \begin{cases} \sup\limits_{\theta\in\Theta}(\mathrm{Cr}\{\theta\}\wedge\mathrm{Pr}\{\Lambda^c(\theta)\}), & \text{if } \sup\limits_{\theta\in\Theta}(\mathrm{Cr}\{\theta\}\wedge\mathrm{Pr}\{\Lambda^c(\theta)\}) < 0.5 \\ 1 - \sup\limits_{\theta\in\Theta}(\mathrm{Cr}\{\theta\}\wedge\mathrm{Pr}\{\Lambda(\theta)\}), & \text{if } \sup\limits_{\theta\in\Theta}(\mathrm{Cr}\{\theta\}\wedge\mathrm{Pr}\{\Lambda^c(\theta)\}) \geq 0.5. \end{cases}$$

The argument breaks down into three cases.

Case 1: $\sup\limits_{\theta\in\Theta}(\mathrm{Cr}\{\theta\}\wedge\mathrm{Pr}\{\Lambda(\theta)\}) < 0.5$. For this case, we have

$$\sup_{\theta\in\Theta}(\mathrm{Cr}\{\theta\}\wedge\mathrm{Pr}\{\Lambda^c(\theta)\}) \geq 0.5,$$

$$\mathrm{Ch}\{\Lambda\} + \mathrm{Ch}\{\Lambda^c\} = \sup_{\theta\in\Theta}(\mathrm{Cr}\{\theta\}\wedge\mathrm{Pr}\{\Lambda(\theta)\})+1-\sup_{\theta\in\Theta}(\mathrm{Cr}\{\theta\}\wedge\mathrm{Pr}\{\Lambda(\theta)\}) = 1.$$

Case 2: $\sup\limits_{\theta\in\Theta}(\mathrm{Cr}\{\theta\}\wedge\mathrm{Pr}\{\Lambda(\theta)\}) \geq 0.5$ and $\sup\limits_{\theta\in\Theta}(\mathrm{Cr}\{\theta\}\wedge\mathrm{Pr}\{\Lambda^c(\theta)\}) < 0.5$. For this case, we have

$$\mathrm{Ch}\{\Lambda\}+\mathrm{Ch}\{\Lambda^c\} = 1-\sup_{\theta\in\Theta}(\mathrm{Cr}\{\theta\}\wedge\mathrm{Pr}\{\Lambda^c(\theta)\})+\sup_{\theta\in\Theta}(\mathrm{Cr}\{\theta\}\wedge\mathrm{Pr}\{\Lambda^c(\theta)\}) = 1.$$

Case 3: $\sup\limits_{\theta\in\Theta}(\mathrm{Cr}\{\theta\}\wedge\mathrm{Pr}\{\Lambda(\theta)\}) \geq 0.5$ and $\sup\limits_{\theta\in\Theta}(\mathrm{Cr}\{\theta\}\wedge\mathrm{Pr}\{\Lambda^c(\theta)\}) \geq 0.5$. For this case, it follows from Theorem C.3 that

$$\sup_{\theta\in\Theta}(\mathrm{Cr}\{\theta\}\wedge\mathrm{Pr}\{\Lambda(\theta)\}) = \sup_{\theta\in\Theta}(\mathrm{Cr}\{\theta\}\wedge\mathrm{Pr}\{\Lambda^c(\theta)\}) = 0.5.$$

Hence $\mathrm{Ch}\{\Lambda\} + \mathrm{Ch}\{\Lambda^c\} = 0.5 + 0.5 = 1$. The theorem is proved.

Theorem C.6 *(Li and Liu [98]) For any event $X \times Y$, we have*

$$\text{Ch}\{X \times Y\} = \text{Cr}\{X\} \wedge \text{Pr}\{Y\}. \tag{C.11}$$

Proof: The argument breaks down into three cases.

Case 1: $\text{Cr}\{X\} < 0.5$. For this case, we have

$$\sup_{\theta \in X} \text{Cr}\{\theta\} \wedge \text{Pr}\{Y\} = \text{Cr}\{X\} \wedge \text{Cr}\{Y\} < 0.5,$$

$$\text{Ch}\{X \times Y\} = \sup_{\theta \in X} \text{Cr}\{\theta\} \wedge \text{Pr}\{Y\} = \text{Cr}\{X\} \wedge \text{Pr}\{Y\}.$$

Case 2: $\text{Cr}\{X\} \geq 0.5$ and $\text{Pr}\{Y\} < 0.5$. Then we have

$$\sup_{\theta \in X} \text{Cr}\{\theta\} \geq 0.5,$$

$$\sup_{\theta \in X} \text{Cr}\{\theta\} \wedge \text{Pr}\{Y\} = \text{Pr}\{Y\} < 0.5,$$

$$\text{Ch}\{X \times Y\} = \sup_{\theta \in X} \text{Cr}\{\theta\} \wedge \text{Pr}\{Y\} = \text{Pr}\{Y\} = \text{Cr}\{X\} \wedge \text{Pr}\{Y\}.$$

Case 3: $\text{Cr}\{X\} \geq 0.5$ and $\text{Pr}\{Y\} \geq 0.5$. Then we have

$$\sup_{\theta \in \Theta} (\text{Cr}\{\theta\} \wedge \text{Pr}\{(X \times Y)(\theta)\}) \geq \sup_{\theta \in X} \text{Cr}\{\theta\} \wedge \text{Pr}\{Y\} \geq 0.5,$$

$$\text{Ch}\{X \times Y\} = 1 - \sup_{\theta \in \Theta} (\text{Cr}\{\theta\} \wedge \text{Pr}\{(X \times Y)^c(\theta)\}) = \text{Cr}\{X\} \wedge \text{Pr}\{Y\}.$$

The theorem is proved.

Example C.4: It follows from Theorem C.6 that for any events $X \times \Omega$ and $\Theta \times Y$, we have

$$\text{Ch}\{X \times \Omega\} = \text{Cr}\{X\}, \quad \text{Ch}\{\Theta \times Y\} = \text{Pr}\{Y\}. \tag{C.12}$$

Theorem C.7 *(Li and Liu [98], Chance Subadditivity Theorem) The chance measure is subadditive. That is,*

$$\text{Ch}\{\Lambda_1 \cup \Lambda_2\} \leq \text{Ch}\{\Lambda_1\} + \text{Ch}\{\Lambda_2\} \tag{C.13}$$

for any events Λ_1 and Λ_2. In fact, chance measure is not only finitely subadditive but also countably subadditive.

Proof: The proof breaks down into three cases.

Case 1: $\text{Ch}\{\Lambda_1 \cup \Lambda_2\} < 0.5$. Then $\text{Ch}\{\Lambda_1\} < 0.5$, $\text{Ch}\{\Lambda_2\} < 0.5$ and

$$\begin{aligned}
\text{Ch}\{\Lambda_1 \cup \Lambda_2\} &= \sup_{\theta \in \Theta} (\text{Cr}\{\theta\} \wedge \text{Pr}\{(\Lambda_1 \cup \Lambda_2)(\theta)\}) \\
&\leq \sup_{\theta \in \Theta} (\text{Cr}\{\theta\} \wedge (\text{Pr}\{\Lambda_1(\theta)\} + \text{Pr}\{\Lambda_2(\theta)\})) \\
&\leq \sup_{\theta \in \Theta} (\text{Cr}\{\theta\} \wedge \text{Pr}\{\Lambda_1(\theta)\} + \text{Cr}\{\theta\} \wedge \text{Pr}\{\Lambda_2(\theta)\}) \\
&\leq \sup_{\theta \in \Theta} (\text{Cr}\{\theta\} \wedge \text{Pr}\{\Lambda_1(\theta)\}) + \sup_{\theta \in \Theta} (\text{Cr}\{\theta\} \wedge \text{Pr}\{\Lambda_2(\theta)\}) \\
&= \text{Ch}\{\Lambda_1\} + \text{Ch}\{\Lambda_2\}.
\end{aligned}$$

Case 2: $\mathrm{Ch}\{\Lambda_1 \cup \Lambda_2\} \geq 0.5$ and $\mathrm{Ch}\{\Lambda_1\} \vee \mathrm{Ch}\{\Lambda_2\} < 0.5$. We first have

$$\sup_{\theta \in \Theta}(\mathrm{Cr}\{\theta\} \wedge \mathrm{Pr}\{(\Lambda_1 \cup \Lambda_2)(\theta)\}) \geq 0.5.$$

For any sufficiently small number $\varepsilon > 0$, there exists a point θ such that

$$\mathrm{Cr}\{\theta\} \wedge \mathrm{Pr}\{(\Lambda_1 \cup \Lambda_2)(\theta)\} > 0.5 - \varepsilon > \mathrm{Ch}\{\Lambda_1\} \vee \mathrm{Ch}\{\Lambda_2\},$$

$$\mathrm{Cr}\{\theta\} > 0.5 - \varepsilon > \mathrm{Pr}\{\Lambda_1(\theta)\},$$
$$\mathrm{Cr}\{\theta\} > 0.5 - \varepsilon > \mathrm{Pr}\{\Lambda_2(\theta)\}.$$

Thus we have

$$\begin{aligned}
&\mathrm{Cr}\{\theta\} \wedge \mathrm{Pr}\{(\Lambda_1 \cup \Lambda_2)^c(\theta)\} + \mathrm{Cr}\{\theta\} \wedge \mathrm{Pr}\{\Lambda_1(\theta)\} + \mathrm{Cr}\{\theta\} \wedge \mathrm{Pr}\{\Lambda_2(\theta)\} \\
&= \mathrm{Cr}\{\theta\} \wedge \mathrm{Pr}\{(\Lambda_1 \cup \Lambda_2)^c(\theta)\} + \mathrm{Pr}\{\Lambda_1(\theta)\} + \mathrm{Pr}\{\Lambda_2(\theta)\} \\
&\geq \mathrm{Cr}\{\theta\} \wedge \mathrm{Pr}\{(\Lambda_1 \cup \Lambda_2)^c(\theta)\} + \mathrm{Pr}\{(\Lambda_1 \cup \Lambda_2)(\theta)\} \geq 1 - 2\varepsilon
\end{aligned}$$

because if $\mathrm{Cr}\{\theta\} \geq \mathrm{Pr}\{(\Lambda_1 \cup \Lambda_2)^c(\theta)\}$, then

$$\begin{aligned}
&\mathrm{Cr}\{\theta\} \wedge \mathrm{Pr}\{(\Lambda_1 \cup \Lambda_2)^c(\theta)\} + \mathrm{Pr}\{(\Lambda_1 \cup \Lambda_2)(\theta)\} \\
&= \mathrm{Pr}\{(\Lambda_1 \cup \Lambda_2)^c(\theta)\} + \mathrm{Pr}\{(\Lambda_1 \cup \Lambda_2)(\theta)\} \\
&= 1 \geq 1 - 2\varepsilon
\end{aligned}$$

and if $\mathrm{Cr}\{\theta\} < \mathrm{Pr}\{(\Lambda_1 \cup \Lambda_2)^c(\theta)\}$, then

$$\begin{aligned}
&\mathrm{Cr}\{\theta\} \wedge \mathrm{Pr}\{(\Lambda_1 \cup \Lambda_2)^c(\theta)\} + \mathrm{Pr}\{(\Lambda_1 \cup \Lambda_2)(\theta)\} \\
&= \mathrm{Cr}\{\theta\} + \mathrm{Pr}\{(\Lambda_1 \cup \Lambda_2)(\theta)\} \\
&\geq (0.5 - \varepsilon) + (0.5 - \varepsilon) = 1 - 2\varepsilon.
\end{aligned}$$

Taking supremum on both sides and letting $\varepsilon \to 0$, we obtain

$$\begin{aligned}
\mathrm{Ch}\{\Lambda_1 \cup \Lambda_2\} &= 1 - \sup_{\theta \in \Theta}(\mathrm{Cr}\{\theta\} \wedge \mathrm{Pr}\{(\Lambda_1 \cup \Lambda_2)^c(\theta)\}) \\
&\leq \sup_{\theta \in \Theta}(\mathrm{Cr}\{\theta\} \wedge \mathrm{Pr}\{\Lambda_1(\theta)\}) + \sup_{\theta \in \Theta}(\mathrm{Cr}\{\theta\} \wedge \mathrm{Pr}\{\Lambda_2(\theta)\}) \\
&= \mathrm{Ch}\{\Lambda_1\} + \mathrm{Ch}\{\Lambda_2\}.
\end{aligned}$$

Case 3: $\mathrm{Ch}\{\Lambda_1 \cup \Lambda_2\} \geq 0.5$ and $\mathrm{Ch}\{\Lambda_1\} \vee \mathrm{Ch}\{\Lambda_2\} \geq 0.5$. Without loss of generality, suppose $\mathrm{Ch}\{\Lambda_1\} \geq 0.5$. For each θ, we first have

$$\begin{aligned}
\mathrm{Cr}\{\theta\} \wedge \mathrm{Pr}\{\Lambda_1^c(\theta)\} &= \mathrm{Cr}\{\theta\} \wedge \mathrm{Pr}\{(\Lambda_1^c(\theta) \cap \Lambda_2^c(\theta)) \cup (\Lambda_1^c(\theta) \cap \Lambda_2(\theta))\} \\
&\leq \mathrm{Cr}\{\theta\} \wedge (\mathrm{Pr}\{(\Lambda_1 \cup \Lambda_2)^c(\theta)\} + \mathrm{Pr}\{\Lambda_2(\theta)\}) \\
&\leq \mathrm{Cr}\{\theta\} \wedge \mathrm{Pr}\{(\Lambda_1 \cup \Lambda_2)^c(\theta)\} + \mathrm{Cr}\{\theta\} \wedge \mathrm{Pr}\{\Lambda_2(\theta)\},
\end{aligned}$$

i.e., $\text{Cr}\{\theta\} \wedge \text{Pr}\{(\Lambda_1 \cup \Lambda_2)^c(\theta)\} \geq \text{Cr}\{\theta\} \wedge \text{Pr}\{\Lambda_1^c(\theta)\} - \text{Cr}\{\theta\} \wedge \text{Pr}\{\Lambda_2(\theta)\}$. It follows from Theorem C.3 that

$$
\begin{aligned}
\text{Ch}\{\Lambda_1 \cup \Lambda_2\} &= 1 - \sup_{\theta \in \Theta}(\text{Cr}\{\theta\} \wedge \text{Pr}\{(\Lambda_1 \cup \Lambda_2)^c(\theta)\}) \\
&\leq 1 - \sup_{\theta \in \Theta}(\text{Cr}\{\theta\} \wedge \text{Pr}\{\Lambda_1^c(\theta)\}) + \sup_{\theta \in \Theta}(\text{Cr}\{\theta\} \wedge \text{Pr}\{\Lambda_2(\theta)\}) \\
&\leq \text{Ch}\{\Lambda_1\} + \text{Ch}\{\Lambda_2\}.
\end{aligned}
$$

The theorem is proved.

Remark C.1: For any events Λ_1 and Λ_2, it follows from the chance subadditivity theorem that the chance measure is null-additive, i.e., $\text{Ch}\{\Lambda_1 \cup \Lambda_2\} = \text{Ch}\{\Lambda_1\} + \text{Ch}\{\Lambda_2\}$ if either $\text{Ch}\{\Lambda_1\} = 0$ or $\text{Ch}\{\Lambda_2\} = 0$.

Theorem C.8 *Let $\{\Lambda_i\}$ be a decreasing sequence of events with $\text{Ch}\{\Lambda_i\} \to 0$ as $i \to \infty$. Then for any event Λ, we have*

$$
\lim_{i \to \infty} \text{Ch}\{\Lambda \cup \Lambda_i\} = \lim_{i \to \infty} \text{Ch}\{\Lambda \backslash \Lambda_i\} = \text{Ch}\{\Lambda\}. \tag{C.14}
$$

Proof: Since chance measure is increasing and subadditive, we immediately have

$$
\text{Ch}\{\Lambda\} \leq \text{Ch}\{\Lambda \cup \Lambda_i\} \leq \text{Ch}\{\Lambda\} + \text{Ch}\{\Lambda_i\}
$$

for each i. Thus we get $\text{Ch}\{\Lambda \cup \Lambda_i\} \to \text{Ch}\{\Lambda\}$ by using $\text{Ch}\{\Lambda_i\} \to 0$. Since $(\Lambda \backslash \Lambda_i) \subset \Lambda \subset ((\Lambda \backslash \Lambda_i) \cup \Lambda_i)$, we have

$$
\text{Ch}\{\Lambda \backslash \Lambda_i\} \leq \text{Ch}\{\Lambda\} \leq \text{Ch}\{\Lambda \backslash \Lambda_i\} + \text{Ch}\{\Lambda_i\}.
$$

Hence $\text{Ch}\{\Lambda \backslash \Lambda_i\} \to \text{Ch}\{\Lambda\}$ by using $\text{Ch}\{\Lambda_i\} \to 0$.

Theorem C.9 *(Li and Liu [98], Chance Semicontinuity Law) For events $\Lambda_1, \Lambda_2, \cdots$, we have*

$$
\lim_{i \to \infty} \text{Ch}\{\Lambda_i\} = \text{Ch}\left\{\lim_{i \to \infty} \Lambda_i\right\} \tag{C.15}
$$

if one of the following conditions is satisfied:
(a) $\text{Ch}\{\Lambda\} \leq 0.5$ and $\Lambda_i \uparrow \Lambda$;   *(b) $\lim_{i \to \infty} \text{Ch}\{\Lambda_i\} < 0.5$ and $\Lambda_i \uparrow \Lambda$;*
(c) $\text{Ch}\{\Lambda\} \geq 0.5$ and $\Lambda_i \downarrow \Lambda$;   *(d) $\lim_{i \to \infty} \text{Ch}\{\Lambda_i\} > 0.5$ and $\Lambda_i \downarrow \Lambda$.*

Proof: (a) Assume $\text{Ch}\{\Lambda\} \leq 0.5$ and $\Lambda_i \uparrow \Lambda$. We first have

$$
\text{Ch}\{\Lambda\} = \sup_{\theta \in \Theta}(\text{Cr}\{\theta\} \wedge \text{Pr}\{\Lambda(\theta)\}), \quad \text{Ch}\{\Lambda_i\} = \sup_{\theta \in \Theta}(\text{Cr}\{\theta\} \wedge \text{Pr}\{\Lambda_i(\theta)\})
$$

for $i = 1, 2, \cdots$ For each $\theta \in \Theta$, since $\Lambda_i(\theta) \uparrow \Lambda(\theta)$, it follows from the probability continuity theorem that

$$
\lim_{i \to \infty} \text{Cr}\{\theta\} \wedge \text{Pr}\{\Lambda_i(\theta)\} = \text{Cr}\{\theta\} \wedge \text{Pr}\{\Lambda(\theta)\}.
$$

Taking supremum on both sides, we obtain

$$\lim_{i \to \infty} \sup_{\theta \in \Theta} (\mathrm{Cr}\{\theta\} \wedge \mathrm{Pr}\{\Lambda_i(\theta)\}) = \sup_{\theta \in \Theta} (\mathrm{Cr}\{\theta\} \wedge \mathrm{Pr}\{\Lambda(\theta)\}).$$

The part (a) is verified.

(b) Assume $\lim_{i \to \infty} \mathrm{Ch}\{\Lambda_i\} < 0.5$ and $\Lambda_i \uparrow \Lambda$. For each $\theta \in \Theta$, since

$$\mathrm{Cr}\{\theta\} \wedge \mathrm{Pr}\{\Lambda(\theta)\} = \lim_{i \to \infty} \mathrm{Cr}\{\theta\} \wedge \mathrm{Pr}\{\Lambda_i(\theta)\},$$

we have

$$\sup_{\theta \in \Theta} (\mathrm{Cr}\{\theta\} \wedge \mathrm{Pr}\{\Lambda(\theta)\}) \leq \lim_{i \to \infty} \sup_{\theta \in \Theta} (\mathrm{Cr}\{\theta\} \wedge \mathrm{Pr}\{\Lambda_i(\theta)\}) < 0.5.$$

It follows that $\mathrm{Ch}\{\Lambda\} < 0.5$ and the part (b) holds by using (a).

(c) Assume $\mathrm{Ch}\{\Lambda\} \geq 0.5$ and $\Lambda_i \downarrow \Lambda$. We have $\mathrm{Ch}\{\Lambda^c\} \leq 0.5$ and $\Lambda_i^c \uparrow \Lambda^c$. It follows from (a) that

$$\lim_{i \to \infty} \mathrm{Ch}\{\Lambda_i\} = 1 - \lim_{i \to \infty} \mathrm{Ch}\{\Lambda_i^c\} = 1 - \mathrm{Ch}\{\Lambda^c\} = \mathrm{Ch}\{\Lambda\}.$$

(d) Assume $\lim_{i \to \infty} \mathrm{Ch}\{\Lambda_i\} > 0.5$ and $\Lambda_i \downarrow \Lambda$. We have $\lim_{i \to \infty} \mathrm{Ch}\{\Lambda_i^c\} < 0.5$ and $\Lambda_i^c \uparrow \Lambda^c$. It follows from (b) that

$$\lim_{i \to \infty} \mathrm{Ch}\{\Lambda_i\} = 1 - \lim_{i \to \infty} \mathrm{Ch}\{\Lambda_i^c\} = 1 - \mathrm{Ch}\{\Lambda^c\} = \mathrm{Ch}\{\Lambda\}.$$

The theorem is proved.

Theorem C.10 *(Chance Asymptotic Theorem) For any events $\Lambda_1, \Lambda_2, \cdots$, we have*

$$\lim_{i \to \infty} \mathrm{Ch}\{\Lambda_i\} \geq 0.5, \quad \textit{if } \Lambda_i \uparrow \Theta \times \Omega, \tag{C.16}$$

$$\lim_{i \to \infty} \mathrm{Ch}\{\Lambda_i\} \leq 0.5, \quad \textit{if } \Lambda_i \downarrow \emptyset. \tag{C.17}$$

Proof: Assume $\Lambda_i \uparrow \Theta \times \Omega$. If $\lim_{i \to \infty} \mathrm{Ch}\{\Lambda_i\} < 0.5$, it follows from the chance semicontinuity law that

$$\mathrm{Ch}\{\Theta \times \Omega\} = \lim_{i \to \infty} \mathrm{Ch}\{\Lambda_i\} < 0.5$$

which is in contradiction with $\mathrm{Cr}\{\Theta \times \Omega\} = 1$. The first inequality is proved. The second one may be verified similarly.

C.2   Hybrid Variable

Recall that a random variable is a measurable function from a probability space to the set of real numbers, and a fuzzy variable is a function from a credibility space to the set of real numbers. In order to describe a quantity with both fuzziness and randomness, we introduce a concept of hybrid variable as follows.
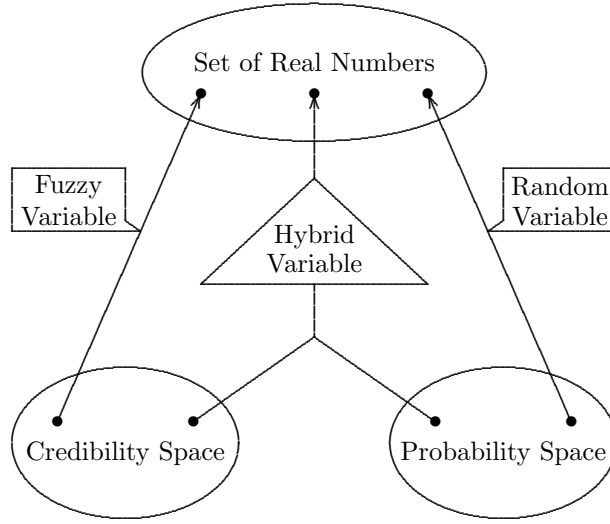
Figure C.2: Graphical Representation of Hybrid Variable

Definition C.4 *(Liu [125]) A hybrid variable is a measurable function from a chance space $(\Theta, \mathcal{P}, \mathrm{Cr}) \times (\Omega, \mathcal{A}, \mathrm{Pr})$ to the set of real numbers, i.e., for any Borel set B of real numbers, the set*

$$\{\xi \in B\} = \{(\theta, \omega) \in \Theta \times \Omega \mid \xi(\theta, \omega) \in B\} \qquad (C.18)$$

is an event.

Remark C.2: A hybrid variable degenerates to a fuzzy variable if the value of $\xi(\theta, \omega)$ does not vary with ω. For example,

$$\xi(\theta, \omega) = \theta, \quad \xi(\theta, \omega) = \theta^2 + 1, \quad \xi(\theta, \omega) = \sin \theta.$$

Remark C.3: A hybrid variable degenerates to a random variable if the value of $\xi(\theta, \omega)$ does not vary with θ. For example,

$$\xi(\theta, \omega) = \omega, \quad \xi(\theta, \omega) = \omega^2 + 1, \quad \xi(\theta, \omega) = \sin \omega.$$

Remark C.4: For each fixed $\theta \in \Theta$, it is clear that the hybrid variable $\xi(\theta, \omega)$ is a measurable function from the probability space $(\Omega, \mathcal{A}, \mathrm{Pr})$ to the set of real numbers. Thus it is a random variable and we will denote it by $\xi(\theta, \cdot)$. Then a hybrid variable $\xi(\theta, \omega)$ may also be regarded as a function from a credibility space $(\Theta, \mathcal{P}, \mathrm{Cr})$ to the set $\{\xi(\theta, \cdot) | \theta \in \Theta\}$ of random variables. Thus ξ is a *random fuzzy variable* defined by Liu [119].

Remark C.5: For each fixed $\omega \in \Omega$, it is clear that the hybrid variable $\xi(\theta, \omega)$ is a function from the credibility space $(\Theta, \mathcal{P}, \mathrm{Cr})$ to the set of real

numbers. Thus it is a fuzzy variable and we will denote it by $\xi(\cdot, \omega)$. Then a hybrid variable $\xi(\theta, \omega)$ may be regarded as a function from a probability space $(\Omega, \mathcal{A}, \mathrm{Pr})$ to the set $\{\xi(\cdot, \omega)|\omega \in \Omega\}$ of fuzzy variables. If $\mathrm{Cr}\{\xi(\cdot, \omega) \in B\}$ is a measurable function of ω for any Borel set B of real numbers, then ξ is a *fuzzy random variable* in the sense of Liu and Liu [142].

Model I

If $\tilde{a}$ is a fuzzy variable and η is a random variable, then the sum $\xi = \tilde{a} + \eta$ is a hybrid variable. The product $\xi = \tilde{a} \cdot \eta$ is also a hybrid variable. Generally speaking, if $f : \Re^2 \to \Re$ is a measurable function, then

$$\xi = f(\tilde{a}, \eta) \tag{C.19}$$

is a hybrid variable. Suppose that $\tilde{a}$ has a membership function μ, and η has a probability density function ϕ. Then for any Borel set B of real numbers, we have

$$\mathrm{Ch}\{f(\tilde{a}, \eta) \in B\} = \begin{cases} \sup_x \left(\dfrac{\mu(x)}{2} \wedge \displaystyle\int_{f(x,y)\in B} \phi(y)\mathrm{d}y \right), \\ \qquad \text{if } \sup_x \left(\dfrac{\mu(x)}{2} \wedge \displaystyle\int_{f(x,y)\in B} \phi(y)\mathrm{d}y \right) < 0.5 \\ 1 - \sup_x \left(\dfrac{\mu(x)}{2} \wedge \displaystyle\int_{f(x,y)\in B^c} \phi(y)\mathrm{d}y \right), \\ \qquad \text{if } \sup_x \left(\dfrac{\mu(x)}{2} \wedge \displaystyle\int_{f(x,y)\in B} \phi(y)\mathrm{d}y \right) \geq 0.5. \end{cases}$$

More generally, let $\tilde{a}_1, \tilde{a}_2, \cdots, \tilde{a}_m$ be fuzzy variables, and let $\eta_1, \eta_2, \cdots, \eta_n$ be random variables. If $f : \Re^{m+n} \to \Re$ is a measurable function, then

$$\xi = f(\tilde{a}_1, \tilde{a}_2, \cdots, \tilde{a}_m; \eta_1, \eta_2, \cdots, \eta_n) \tag{C.20}$$

is a hybrid variable. The chance $\mathrm{Ch}\{f(\tilde{a}_1, \tilde{a}_2, \cdots, \tilde{a}_m; \eta_1, \eta_2, \cdots, \eta_n) \in B\}$ may be calculated in a similar way provided that μ is the joint membership function and ϕ is the joint probability density function.

Model II

Let $\tilde{a}_1, \tilde{a}_2, \cdots, \tilde{a}_m$ be fuzzy variables, and let $p_1, p_2, \cdots, p_m$ be nonnegative numbers with $p_1 + p_2 + \cdots + p_m = 1$. Then

$$\xi = \begin{cases} \tilde{a}_1 \text{ with probability } p_1 \\ \tilde{a}_2 \text{ with probability } p_2 \\ \cdots \\ \tilde{a}_m \text{ with probability } p_m \end{cases} \tag{C.21}$$

is clearly a hybrid variable. If $\tilde{a}_1, \tilde{a}_2, \cdots, \tilde{a}_m$ have membership functions $\mu_1, \mu_2, \cdots, \mu_m$, respectively, then for any set B of real numbers, we have

$$\text{Ch}\{\xi \in B\} = \begin{cases} \sup_{x_1, x_2 \cdots, x_m} \left(\left(\min_{1 \leq i \leq m} \frac{\mu_i(x_i)}{2} \right) \wedge \sum_{i=1}^{m} \{p_i \,|\, x_i \in B\} \right), \\ \qquad \text{if } \sup_{x_1, x_2 \cdots, x_m} \left(\left(\min_{1 \leq i \leq m} \frac{\mu_i(x_i)}{2} \right) \wedge \sum_{i=1}^{m} \{p_i \,|\, x_i \in B\} \right) < 0.5 \\ 1 - \sup_{x_1, x_2 \cdots, x_m} \left(\left(\min_{1 \leq i \leq m} \frac{\mu_i(x_i)}{2} \right) \wedge \sum_{i=1}^{m} \{p_i \,|\, x_i \in B^c\} \right), \\ \qquad \text{if } \sup_{x_1, x_2 \cdots, x_m} \left(\left(\min_{1 \leq i \leq m} \frac{\mu_i(x_i)}{2} \right) \wedge \sum_{i=1}^{m} \{p_i \,|\, x_i \in B\} \right) \geq 0.5. \end{cases}$$

Model III

Let $\eta_1, \eta_2, \cdots, \eta_m$ be random variables, and let $u_1, u_2, \cdots, u_m$ be nonnegative numbers with $u_1 \vee u_2 \vee \cdots \vee u_m = 1$. Then

$$\xi = \begin{cases} \eta_1 \text{ with membership degree } u_1 \\ \eta_2 \text{ with membership degree } u_2 \\ \qquad \cdots \\ \eta_m \text{ with membership degree } u_m \end{cases} \tag{C.22}$$

is clearly a hybrid variable. If $\eta_1, \eta_2, \cdots, \eta_m$ have probability density functions $\phi_1, \phi_2, \cdots, \phi_m$, respectively, then for any Borel set B of real numbers, we have

$$\text{Ch}\{\xi \in B\} = \begin{cases} \max_{1 \leq i \leq m} \left(\frac{u_i}{2} \wedge \int_B \phi_i(x)\mathrm{d}x \right), \\ \qquad \text{if } \max_{1 \leq i \leq m} \left(\frac{u_i}{2} \wedge \int_B \phi_i(x)\mathrm{d}x \right) < 0.5 \\ 1 - \max_{1 \leq i \leq m} \left(\frac{u_i}{2} \wedge \int_{B^c} \phi_i(x)\mathrm{d}x \right), \\ \qquad \text{if } \max_{1 \leq i \leq m} \left(\frac{u_i}{2} \wedge \int_B \phi_i(x)\mathrm{d}x \right) \geq 0.5. \end{cases}$$

Model IV

In many statistics problems, the probability density function is completely known except for the values of one or more parameters. For example, it might be known that the lifetime ξ of a modern engine is an exponentially

distributed random variable with an unknown expected value β. Usually, there is some relevant information in practice. It is thus possible to specify an interval in which the value of β is likely to lie, or to give an approximate estimate of the value of β. It is typically not possible to determine the value of β exactly. If the value of β is provided as a fuzzy variable, then ξ is a hybrid variable. More generally, suppose that ξ has a probability density function

$$\phi(x; \tilde{a}_1, \tilde{a}_2, \cdots, \tilde{a}_m), \quad x \in \Re \qquad (C.23)$$

in which the parameters $\tilde{a}_1, \tilde{a}_2, \cdots, \tilde{a}_m$ are fuzzy variables rather than crisp numbers. Then ξ is a hybrid variable provided that $\phi(x; y_1, y_2, \cdots, y_m)$ is a probability density function for any $(y_1, y_2, \cdots, y_m)$ that $(\tilde{a}_1, \tilde{a}_2, \cdots, \tilde{a}_m)$ may take. If $\tilde{a}_1, \tilde{a}_2, \cdots, \tilde{a}_m$ have membership functions $\mu_1, \mu_2, \cdots, \mu_m$, respectively, then for any Borel set B of real numbers, the chance $\mathrm{Ch}\{\xi \in B\}$ is

$$
\begin{cases}
\displaystyle\sup_{y_1, y_2 \cdots, y_m} \left(\left(\min_{1 \leq i \leq m} \frac{\mu_i(y_i)}{2} \right) \wedge \int_B \phi(x; y_1, y_2, \cdots, y_m) \mathrm{d}x \right), \\[4mm]
\qquad \text{if } \displaystyle\sup_{y_1, y_2, \cdots, y_m} \left(\left(\min_{1 \leq i \leq m} \frac{\mu_i(y_i)}{2} \right) \wedge \int_B \phi(x; y_1, y_2, \cdots, y_m) \mathrm{d}x \right) < 0.5 \\[4mm]
1 - \displaystyle\sup_{y_1, y_2 \cdots, y_m} \left(\left(\min_{1 \leq i \leq m} \frac{\mu_i(y_i)}{2} \right) \wedge \int_{B^c} \phi(x; y_1, y_2, \cdots, y_m) \mathrm{d}x \right), \\[4mm]
\qquad \text{if } \displaystyle\sup_{y_1, y_2, \cdots, y_m} \left(\left(\min_{1 \leq i \leq m} \frac{\mu_i(y_i)}{2} \right) \wedge \int_B \phi(x; y_1, y_2, \cdots, y_m) \mathrm{d}x \right) \geq 0.5.
\end{cases}
$$

When are two hybrid variables equal to each other?

Definition C.5 *Let ξ_1 and ξ_2 be hybrid variables defined on the chance space $(\Theta, \mathcal{P}, \mathrm{Cr}) \times (\Omega, \mathcal{A}, \mathrm{Pr})$. We say $\xi_1 = \xi_2$ if $\xi_1(\theta, \omega) = \xi_2(\theta, \omega)$ for almost all $(\theta, \omega) \in \Theta \times \Omega$.*

Hybrid Vectors

Definition C.6 *An n-dimensional hybrid vector is a measurable function from a chance space $(\Theta, \mathcal{P}, \mathrm{Cr}) \times (\Omega, \mathcal{A}, \mathrm{Pr})$ to the set of n-dimensional real vectors, i.e., for any Borel set B of $\Re^n$, the set*

$$\{\boldsymbol{\xi} \in B\} = \big\{ (\theta, \omega) \in \Theta \times \Omega \,\big|\, \boldsymbol{\xi}(\theta, \omega) \in B \big\} \qquad (C.24)$$

is an event.

Theorem C.11 *The vector $(\xi_1, \xi_2, \cdots, \xi_n)$ is a hybrid vector if and only if $\xi_1, \xi_2, \cdots, \xi_n$ are hybrid variables.*

Proof: Write $\boldsymbol{\xi} = (\xi_1, \xi_2, \cdots, \xi_n)$. Suppose that $\boldsymbol{\xi}$ is a hybrid vector on the chance space $(\Theta, \mathcal{P}, \mathrm{Cr}) \times (\Omega, \mathcal{A}, \mathrm{Pr})$. For any Borel set B of $\Re$, the set $B \times \Re^{n-1}$ is a Borel set of $\Re^n$. Thus the set

$$\{\xi_1 \in B\} = \{\xi_1 \in B, \xi_2 \in \Re, \cdots, \xi_n \in \Re\} = \{\boldsymbol{\xi} \in B \times \Re^{n-1}\}$$

is an event. Hence ξ_1 is a hybrid variable. A similar process may prove that $\xi_2, \xi_3, \cdots, \xi_n$ are hybrid variables. Conversely, suppose that all $\xi_1, \xi_2, \cdots, \xi_n$ are hybrid variables on the chance space $(\Theta, \mathcal{P}, \mathrm{Cr}) \times (\Omega, \mathcal{A}, \mathrm{Pr})$. We define

$$\mathcal{B} = \left\{ B \subset \Re^n \mid \{\boldsymbol{\xi} \in B\} \text{ is an event} \right\}.$$

The vector $\boldsymbol{\xi} = (\xi_1, \xi_2, \cdots, \xi_n)$ is proved to be a hybrid vector if we can prove that $\mathcal{B}$ contains all Borel sets of $\Re^n$. First, the class $\mathcal{B}$ contains all open intervals of $\Re^n$ because

$$\left\{ \boldsymbol{\xi} \in \prod_{i=1}^n (a_i, b_i) \right\} = \bigcap_{i=1}^n \{\xi_i \in (a_i, b_i)\}$$

is an event. Next, the class $\mathcal{B}$ is a σ-algebra of $\Re^n$ because (i) we have $\Re^n \in \mathcal{B}$ since $\{\boldsymbol{\xi} \in \Re^n\} = \Theta \times \Omega$; (ii) if $B \in \mathcal{B}$, then $\{\boldsymbol{\xi} \in B\}$ is an event, and

$$\{\boldsymbol{\xi} \in B^c\} = \{\boldsymbol{\xi} \in B\}^c$$

is an event. This means that $B^c \in \mathcal{B}$; (iii) if $B_i \in \mathcal{B}$ for $i = 1, 2, \cdots$, then $\{\boldsymbol{\xi} \in B_i\}$ are events and

$$\left\{ \boldsymbol{\xi} \in \bigcup_{i=1}^\infty B_i \right\} = \bigcup_{i=1}^\infty \{\boldsymbol{\xi} \in B_i\}$$

is an event. This means that $\cup_i B_i \in \mathcal{B}$. Since the smallest σ-algebra containing all open intervals of $\Re^n$ is just the Borel algebra of $\Re^n$, the class $\mathcal{B}$ contains all Borel sets of $\Re^n$. The theorem is proved.

Hybrid Arithmetic

Definition C.7 *Let $f : \Re^n \to \Re$ be a measurable function, and $\xi_1, \xi_2, \cdots, \xi_n$ hybrid variables on the chance space $(\Theta, \mathcal{P}, \mathrm{Cr}) \times (\Omega, \mathcal{A}, \mathrm{Pr})$. Then $\xi = f(\xi_1, \xi_2, \cdots, \xi_n)$ is a hybrid variable defined as*

$$\xi(\theta, \omega) = f(\xi_1(\theta, \omega), \xi_2(\theta, \omega), \cdots, \xi_n(\theta, \omega)), \quad \forall (\theta, \omega) \in \Theta \times \Omega. \qquad (C.25)$$

Example C.5: Let ξ_1 and ξ_2 be two hybrid variables. Then the sum $\xi = \xi_1 + \xi_2$ is a hybrid variable defined by

$$\xi(\theta, \omega) = \xi_1(\theta, \omega) + \xi_2(\theta, \omega), \quad \forall (\theta, \omega) \in \Theta \times \Omega.$$

The product $\xi = \xi_1 \xi_2$ is also a hybrid variable defined by

$$\xi(\theta, \omega) = \xi_1(\theta, \omega) \cdot \xi_2(\theta, \omega), \quad \forall (\theta, \omega) \in \Theta \times \Omega.$$

Theorem C.12 *Let $\boldsymbol{\xi}$ be an n-dimensional hybrid vector, and $f : \Re^n \to \Re$ a measurable function. Then $f(\boldsymbol{\xi})$ is a hybrid variable.*

Proof: Assume that $\boldsymbol{\xi}$ is a hybrid vector on the chance space $(\Theta, \mathcal{P}, \mathrm{Cr}) \times (\Omega, \mathcal{A}, \mathrm{Pr})$. For any Borel set B of $\Re$, since f is a measurable function, the $f^{-1}(B)$ is a Borel set of $\Re^n$. Thus the set

$$\{f(\boldsymbol{\xi}) \in B\} = \left\{\boldsymbol{\xi} \in f^{-1}(B)\right\}$$

is an event for any Borel set B. Hence $f(\boldsymbol{\xi})$ is a hybrid variable.

C.3 Chance Distribution

Chance distribution has been defined in several ways. Here we accept the following definition of chance distribution of hybrid variables.

Definition C.8 *(Li and Liu [98]) The chance distribution $\Phi \colon \Re \to [0,1]$ of a hybrid variable ξ is defined by*

$$\Phi(x) = \mathrm{Ch}\left\{\xi \le x\right\}. \tag{C.26}$$

Example C.6: Let η be a random variable on a probability space $(\Omega, \mathcal{A}, \mathrm{Pr})$. It is clear that η may be regarded as a hybrid variable on the chance space $(\Theta, \mathcal{P}, \mathrm{Cr}) \times (\Omega, \mathcal{A}, \mathrm{Pr})$ as follows,

$$\xi(\theta, \omega) = \eta(\omega), \quad \forall (\theta, \omega) \in \Theta \times \Omega.$$

Thus its chance distribution is

$$\Phi(x) = \mathrm{Ch}\{\xi \le x\} = \mathrm{Ch}\{\Theta \times \{\eta \le x\}\} = \mathrm{Cr}\{\Theta\} \wedge \mathrm{Pr}\{\eta \le x\} = \mathrm{Pr}\{\eta \le x\}$$

which is just the probability distribution of the random variable η.

Example C.7: Let $\tilde{a}$ be a fuzzy variable on a credibility space $(\Theta, \mathcal{P}, \mathrm{Cr})$. It is clear that $\tilde{a}$ may be regarded as a hybrid variable on the chance space $(\Theta, \mathcal{P}, \mathrm{Cr}) \times (\Omega, \mathcal{A}, \mathrm{Pr})$ as follows,

$$\xi(\theta, \omega) = \tilde{a}(\theta), \quad \forall (\theta, \omega) \in \Theta \times \Omega.$$

Thus its chance distribution is

$$\Phi(x) = \mathrm{Ch}\{\xi \le x\} = \mathrm{Ch}\{\{\tilde{a} \le x\} \times \Omega\} = \mathrm{Cr}\{\tilde{a} \le x\} \wedge \mathrm{Pr}\{\Omega\} = \mathrm{Cr}\{\tilde{a} \le x\}$$

which is just the credibility distribution of the fuzzy variable $\tilde{a}$.

Theorem C.13 *(Sufficient and Necessary Condition for Chance Distribution) A function $\Phi : \Re \to [0,1]$ is a chance distribution if and only if it is an increasing function with*

$$\lim_{x \to -\infty} \Phi(x) \leq 0.5 \leq \lim_{x \to +\infty} \Phi(x), \qquad (C.27)$$

$$\lim_{y \downarrow x} \Phi(y) = \Phi(x) \ if \ \lim_{y \downarrow x} \Phi(y) > 0.5 \ or \ \Phi(x) \geq 0.5. \qquad (C.28)$$

Proof: It is obvious that a chance distribution Φ is an increasing function. The inequalities (C.27) follow from the chance asymptotic theorem immediately. Assume that x is a point at which $\lim_{y \downarrow x} \Phi(y) > 0.5$. That is,

$$\lim_{y \downarrow x} \mathrm{Ch}\{\xi \leq y\} > 0.5.$$

Since $\{\xi \leq y\} \downarrow \{\xi \leq x\}$ as $y \downarrow x$, it follows from the chance semicontinuity law that

$$\Phi(y) = \mathrm{Ch}\{\xi \leq y\} \downarrow \mathrm{Ch}\{\xi \leq x\} = \Phi(x)$$

as $y \downarrow x$. When x is a point at which $\Phi(x) \geq 0.5$, if $\lim_{y \downarrow x} \Phi(y) \neq \Phi(x)$, then we have

$$\lim_{y \downarrow x} \Phi(y) > \Phi(x) \geq 0.5.$$

For this case, we have proved that $\lim_{y \downarrow x} \Phi(y) = \Phi(x)$. Thus (C.28) is proved.

Conversely, suppose $\Phi : \Re \to [0,1]$ is an increasing function satisfying (C.27) and (C.28). Theorem B.17 states that there is a fuzzy variable whose credibility distribution is just $\Phi(x)$. Since a fuzzy variable is a special hybrid variable, the theorem is proved.

C.4   Expected Value

Expected value has been defined in several ways. This book uses the following definition of expected value operator of hybrid variables.

Definition C.9 *(Li and Liu [98]) Let ξ be a hybrid variable. Then the expected value of ξ is defined by*

$$E[\xi] = \int_0^{+\infty} \mathrm{Ch}\{\xi \geq r\}\mathrm{d}r - \int_{-\infty}^0 \mathrm{Ch}\{\xi \leq r\}\mathrm{d}r \qquad (C.29)$$

provided that at least one of the two integrals is finite.

Example C.8: If a hybrid variable ξ degenerates to a random variable η, then

$$\mathrm{Ch}\{\xi \leq x\} = \mathrm{Pr}\{\eta \leq x\}, \quad \mathrm{Ch}\{\xi \geq x\} = \mathrm{Pr}\{\eta \geq x\}, \quad \forall x \in \Re.$$

It follows from (C.29) that $E[\xi] = E[\eta]$. In other words, the expected value operator of hybrid variable coincides with that of random variable.

Example C.9: If a hybrid variable ξ degenerates to a fuzzy variable $\tilde{a}$, then

$$\mathrm{Ch}\{\xi \leq x\} = \mathrm{Cr}\{\tilde{a} \leq x\}, \quad \mathrm{Ch}\{\xi \geq x\} = \mathrm{Cr}\{\tilde{a} \geq x\}, \quad \forall x \in \Re.$$

It follows from (C.29) that $E[\xi] = E[\tilde{a}]$. In other words, the expected value operator of hybrid variable coincides with that of fuzzy variable.

Example C.10: Let $\tilde{a}$ be a fuzzy variable and η a random variable with finite expected values. Then the hybrid variable $\xi = \tilde{a} + \eta$ has expected value $E[\xi] = E[\tilde{a}] + E[\eta]$.

Theorem C.14 *Let ξ be a hybrid variable with finite expected values. Then for any real numbers a and b, we have*

$$E[a\xi + b] = aE[\xi] + b. \tag{C.30}$$

Proof: STEP 1: We first prove that $E[\xi + b] = E[\xi] + b$ for any real number b. If $b \geq 0$, we have

$$
\begin{aligned}
E[\xi + b] &= \int_0^{+\infty} \mathrm{Ch}\{\xi + b \geq r\}\mathrm{d}r - \int_{-\infty}^0 \mathrm{Ch}\{\xi + b \leq r\}\mathrm{d}r \\
&= \int_0^{+\infty} \mathrm{Ch}\{\xi \geq r - b\}\mathrm{d}r - \int_{-\infty}^0 \mathrm{Ch}\{\xi \leq r - b\}\mathrm{d}r \\
&= E[\xi] + \int_0^b (\mathrm{Ch}\{\xi \geq r - b\} + \mathrm{Ch}\{\xi < r - b\})\mathrm{d}r \\
&= E[\xi] + b.
\end{aligned}
$$

If $b < 0$, then we have

$$E[a\xi + b] = E[\xi] - \int_b^0 (\mathrm{Ch}\{\xi \geq r - b\} + \mathrm{Ch}\{\xi < r - b\})\mathrm{d}r = E[\xi] + b.$$

STEP 2: We prove $E[a\xi] = aE[\xi]$. If $a = 0$, then the equation $E[a\xi] = aE[\xi]$ holds trivially. If $a > 0$, we have

$$
\begin{aligned}
E[a\xi] &= \int_0^{+\infty} \mathrm{Ch}\{a\xi \geq r\}\mathrm{d}r - \int_{-\infty}^0 \mathrm{Ch}\{a\xi \leq r\}\mathrm{d}r \\
&= \int_0^{+\infty} \mathrm{Ch}\{\xi \geq r/a\}\mathrm{d}r - \int_{-\infty}^0 \mathrm{Ch}\{\xi \leq r/a\}\mathrm{d}r \\
&= a\int_0^{+\infty} \mathrm{Ch}\{\xi \geq t\}\mathrm{d}t - a\int_{-\infty}^0 \mathrm{Ch}\{\xi \leq t\}\mathrm{d}t \\
&= aE[\xi].
\end{aligned}
$$

If $a < 0$, we have

$$
\begin{aligned}
E[a\xi] &= \int_0^{+\infty} \mathrm{Ch}\{a\xi \geq r\}\mathrm{d}r - \int_{-\infty}^0 \mathrm{Ch}\{a\xi \leq r\}\mathrm{d}r \\
&= \int_0^{+\infty} \mathrm{Ch}\{\xi \leq r/a\}\mathrm{d}r - \int_{-\infty}^0 \mathrm{Ch}\{\xi \geq r/a\}\mathrm{d}r \\
&= a\int_0^{+\infty} \mathrm{Ch}\{\xi \geq t\}\mathrm{d}t - a\int_{-\infty}^0 \mathrm{Ch}\{\xi \leq t\}\mathrm{d}t \\
&= aE[\xi].
\end{aligned}
$$

STEP 3: For any real numbers a and b, it follows from Steps 1 and 2 that

$$
E[a\xi + b] = E[a\xi] + b = aE[\xi] + b.
$$

The theorem is proved.

C.5   Variance

Definition C.10 *(Li and Liu [98]) Let ξ be a hybrid variable with finite expected value e. Then the variance of ξ is defined by $V[\xi] = E[(\xi - e)^2]$.*

Theorem C.15 *If ξ is a hybrid variable with finite expected value, a and b are real numbers, then $V[a\xi + b] = a^2 V[\xi]$.*

Proof: It follows from the definition of variance that

$$
V[a\xi + b] = E\left[(a\xi + b - aE[\xi] - b)^2\right] = a^2 E[(\xi - E[\xi])^2] = a^2 V[\xi].
$$

Theorem C.16 *Let ξ be a hybrid variable with expected value e.   Then $V[\xi] = 0$ if and only if $\mathrm{Ch}\{\xi = e\} = 1$.*

Proof: If $V[\xi] = 0$, then $E[(\xi - e)^2] = 0$. Note that

$$
E[(\xi - e)^2] = \int_0^{+\infty} \mathrm{Ch}\{(\xi - e)^2 \geq r\}\mathrm{d}r
$$

which implies $\mathrm{Ch}\{(\xi - e)^2 \geq r\} = 0$ for any $r > 0$. Hence we have

$$
\mathrm{Ch}\{(\xi - e)^2 = 0\} = 1.
$$

That is, $\mathrm{Ch}\{\xi = e\} = 1$.   Conversely, if $\mathrm{Ch}\{\xi = e\} = 1$, then we have $\mathrm{Ch}\{(\xi - e)^2 = 0\} = 1$ and $\mathrm{Ch}\{(\xi - e)^2 \geq r\} = 0$ for any $r > 0$. Thus

$$
V[\xi] = \int_0^{+\infty} \mathrm{Ch}\{(\xi - e)^2 \geq r\}\mathrm{d}r = 0.
$$

The theorem is proved.

C.6   Critical Values

In order to rank hybrid variables, we introduce the following definition of critical values of hybrid variables.

Definition C.11 *(Li and Liu [98]) Let ξ be a hybrid variable, and $\alpha \in (0, 1]$. Then*

$$\xi_{\sup}(\alpha) = \sup \left\{ r \mid \mathrm{Ch}\left\{\xi \geq r\right\} \geq \alpha \right\} \tag{C.31}$$

is called the α-optimistic value to ξ, and

$$\xi_{\inf}(\alpha) = \inf \left\{ r \mid \mathrm{Ch}\left\{\xi \leq r\right\} \geq \alpha \right\} \tag{C.32}$$

is called the α-pessimistic value to ξ.

The hybrid variable ξ reaches upwards of the α-optimistic value $\xi_{\sup}(\alpha)$, and is below the α-pessimistic value $\xi_{\inf}(\alpha)$ with chance α.

Example C.11: If a hybrid variable ξ degenerates to a random variable η, then

$$\mathrm{Ch}\{\xi \leq x\} = \mathrm{Pr}\{\eta \leq x\}, \quad \mathrm{Ch}\{\xi \geq x\} = \mathrm{Pr}\{\eta \geq x\}, \quad \forall x \in \Re.$$

It follows from the definition of critical values that

$$\xi_{\sup}(\alpha) = \eta_{\sup}(\alpha), \quad \xi_{\inf}(\alpha) = \eta_{\inf}(\alpha), \quad \forall \alpha \in (0, 1].$$

In other words, the critical values of hybrid variable coincide with that of random variable.

Example C.12: If a hybrid variable ξ degenerates to a fuzzy variable $\tilde{a}$, then

$$\mathrm{Ch}\{\xi \leq x\} = \mathrm{Cr}\{\tilde{a} \leq x\}, \quad \mathrm{Ch}\{\xi \geq x\} = \mathrm{Cr}\{\tilde{a} \geq x\}, \quad \forall x \in \Re.$$

It follows from the definition of critical values that

$$\xi_{\sup}(\alpha) = \tilde{a}_{\sup}(\alpha), \quad \xi_{\inf}(\alpha) = \tilde{a}_{\inf}(\alpha), \quad \forall \alpha \in (0, 1].$$

In other words, the critical values of hybrid variable coincide with that of fuzzy variable.

Theorem C.17 *Let ξ be a hybrid variable, and $\alpha \in (0, 1]$. If $\alpha > 0.5$, then we have*

$$\mathrm{Ch}\{\xi \leq \xi_{\inf}(\alpha)\} \geq \alpha, \quad \mathrm{Ch}\{\xi \geq \xi_{\sup}(\alpha)\} \geq \alpha. \tag{C.33}$$

Proof: It follows from the definition of α-pessimistic value that there exists a decreasing sequence $\{x_i\}$ such that $\mathrm{Ch}\{\xi \leq x_i\} \geq \alpha$ and $x_i \downarrow \xi_{\inf}(\alpha)$ as

$i \to \infty$. Since $\{\xi \leq x_i\} \downarrow \{\xi \leq \xi_{\inf}(\alpha)\}$ and $\lim_{i\to\infty} \mathrm{Ch}\{\xi \leq x_i\} \geq \alpha > 0.5$, it follows from the chance semicontinuity theorem that

$$\mathrm{Ch}\{\xi \leq \xi_{\inf}(\alpha)\} = \lim_{i\to\infty} \mathrm{Ch}\{\xi \leq x_i\} \geq \alpha.$$

Similarly, there exists an increasing sequence $\{x_i\}$ such that $\mathrm{Ch}\{\xi \geq x_i\} \geq \alpha$ and $x_i \uparrow \xi_{\sup}(\alpha)$ as $i \to \infty$. Since $\{\xi \geq x_i\} \downarrow \{\xi \geq \xi_{\sup}(\alpha)\}$ and $\lim_{i\to\infty} \mathrm{Ch}\{\xi \geq x_i\} \geq \alpha > 0.5$, it follows from the chance semicontinuity theorem that

$$\mathrm{Ch}\{\xi \geq \xi_{\sup}(\alpha)\} = \lim_{i\to\infty} \mathrm{Ch}\{\xi \geq x_i\} \geq \alpha.$$

The theorem is proved.

Theorem C.18 *Let ξ be a hybrid variable, and $\alpha \in (0, 1]$. Then we have*
(a) $\xi_{\sup}(\alpha)$ is a decreasing and left-continuous function of α;
(b) $\xi_{\inf}(\alpha)$ is an increasing and left-continuous function of α.

Proof: (a) Let α_1 and α_2 be two numbers with $0 < \alpha_1 < \alpha_2 \leq 1$. Then for any number $r < \xi_{\sup}(\alpha_2)$, we have

$$\mathrm{Ch}\{\xi \geq r\} \geq \alpha_2 > \alpha_1.$$

Thus, by the definition of optimistic value, we obtain $\xi_{\sup}(\alpha_1) \geq \xi_{\sup}(\alpha_2)$. That is, the value $\xi_{\sup}(\alpha)$ is a decreasing function of α. Next, we prove the left-continuity of $\xi_{\inf}(\alpha)$ with respect to α. Let $\{\alpha_i\}$ be an arbitrary sequence of positive numbers such that $\alpha_i \uparrow \alpha$. Then $\{\xi_{\inf}(\alpha_i)\}$ is an increasing sequence. If the limitation is equal to $\xi_{\inf}(\alpha)$, then the left-continuity is proved. Otherwise, there exists a number z^* such that

$$\lim_{i\to\infty} \xi_{\inf}(\alpha_i) < z^* < \xi_{\inf}(\alpha).$$

Thus $\mathrm{Ch}\{\xi \leq z^*\} \geq \alpha_i$ for each i. Letting $i \to \infty$, we get $\mathrm{Ch}\{\xi \leq z^*\} \geq \alpha$. Hence $z^* \geq \xi_{\inf}(\alpha)$. A contradiction proves the left-continuity of $\xi_{\inf}(\alpha)$ with respect to α. The part (b) may be proved similarly.

C.7   Conditional Chance

We consider the chance measure of an event A after it has been learned that some other event B has occurred. This new chance measure of A is called the conditional chance measure of A given B.

In order to define a conditional chance measure $\mathrm{Ch}\{A|B\}$, at first we have to enlarge $\mathrm{Ch}\{A \cap B\}$ because $\mathrm{Ch}\{A \cap B\} < 1$ for all events whenever $\mathrm{Ch}\{B\} < 1$. It seems that we have no alternative but to divide $\mathrm{Ch}\{A \cap B\}$ by $\mathrm{Ch}\{B\}$. Unfortunately, $\mathrm{Ch}\{A \cap B\}/\mathrm{Ch}\{B\}$ is not always a chance measure.

However, the value $\text{Ch}\{A|B\}$ should not be greater than $\text{Ch}\{A \cap B\}/\text{Ch}\{B\}$ (otherwise the normality will be lost), i.e.,

$$\text{Ch}\{A|B\} \le \frac{\text{Ch}\{A \cap B\}}{\text{Ch}\{B\}}. \tag{C.34}$$

On the other hand, in order to preserve the self-duality, we should have

$$\text{Ch}\{A|B\} = 1 - \text{Ch}\{A^c|B\} \ge 1 - \frac{\text{Ch}\{A^c \cap B\}}{\text{Ch}\{B\}}. \tag{C.35}$$

Furthermore, since $(A \cap B) \cup (A^c \cap B) = B$, we have $\text{Ch}\{B\} \le \text{Ch}\{A \cap B\} + \text{Ch}\{A^c \cap B\}$ by using the chance subadditivity theorem. Thus

$$0 \le 1 - \frac{\text{Ch}\{A^c \cap B\}}{\text{Ch}\{B\}} \le \frac{\text{Ch}\{A \cap B\}}{\text{Ch}\{B\}} \le 1. \tag{C.36}$$

Hence any numbers between $1 - \text{Ch}\{A^c \cap B\}/\text{Ch}\{B\}$ and $\text{Ch}\{A \cap B\}/\text{Ch}\{B\}$ are reasonable values that the conditional chance may take. Based on the maximum uncertainty principle, we have the following conditional chance measure.

Definition C.12 *(Li and Liu [101]) Let $(\Theta, \mathcal{P}, \text{Cr}) \times (\Omega, \mathcal{A}, \text{Pr})$ be a chance space and A, B two events. Then the conditional chance measure of A given B is defined by*

$$\text{Ch}\{A|B\} = \begin{cases} \dfrac{\text{Ch}\{A \cap B\}}{\text{Ch}\{B\}}, & \text{if } \dfrac{\text{Ch}\{A \cap B\}}{\text{Ch}\{B\}} < 0.5 \\[2ex] 1 - \dfrac{\text{Ch}\{A^c \cap B\}}{\text{Ch}\{B\}}, & \text{if } \dfrac{\text{Ch}\{A^c \cap B\}}{\text{Ch}\{B\}} < 0.5 \\[2ex] 0.5, & \text{otherwise} \end{cases} \tag{C.37}$$

provided that $\text{Ch}\{B\} > 0$.

Remark C.6: It follows immediately from the definition of conditional chance that

$$1 - \frac{\text{Ch}\{A^c \cap B\}}{\text{Ch}\{B\}} \le \text{Ch}\{A|B\} \le \frac{\text{Ch}\{A \cap B\}}{\text{Ch}\{B\}}. \tag{C.38}$$

Furthermore, it is clear that the conditional chance measure obeys the maximum uncertainty principle.

Remark C.7: Let X and Y be events in the credibility space. Then the conditional chance measure of $X \times \Omega$ given $Y \times \Omega$ is

$$\text{Ch}\{X \times \Omega | Y \times \Omega\} = \begin{cases} \dfrac{\text{Cr}\{X \cap Y\}}{\text{Cr}\{Y\}}, & \text{if } \dfrac{\text{Cr}\{X \cap Y\}}{\text{Cr}\{Y\}} < 0.5 \\[2ex] 1 - \dfrac{\text{Cr}\{X^c \cap Y\}}{\text{Cr}\{Y\}}, & \text{if } \dfrac{\text{Cr}\{X^c \cap Y\}}{\text{Cr}\{Y\}} < 0.5 \\[2ex] 0.5, & \text{otherwise} \end{cases}$$

which is just the conditional credibility of X given Y.

Remark C.8: Let X and Y be events in the probability space. Then the conditional chance measure of $\Theta \times X$ given $\Theta \times Y$ is

$$\text{Ch}\{\Theta \times X | \Theta \times Y\} = \frac{\text{Pr}\{X \cap Y\}}{\text{Pr}\{Y\}}$$

which is just the conditional probability of X given Y.

Theorem C.19 *(Li and Liu [101]) Conditional chance measure is normal, increasing, self-dual and countably subadditive.*

Proof: At first, the conditional chance measure $\text{Ch}\{\cdot | B\}$ is normal, i.e.,

$$\text{Ch}\{\Theta \times \Omega | B\} = 1 - \frac{\text{Ch}\{\emptyset\}}{\text{Ch}\{B\}} = 1.$$

For any events A_1 and A_2 with $A_1 \subset A_2$, if

$$\frac{\text{Ch}\{A_1 \cap B\}}{\text{Ch}\{B\}} \leq \frac{\text{Ch}\{A_2 \cap B\}}{\text{Ch}\{B\}} < 0.5,$$

then

$$\text{Ch}\{A_1 | B\} = \frac{\text{Ch}\{A_1 \cap B\}}{\text{Ch}\{B\}} \leq \frac{\text{Ch}\{A_2 \cap B\}}{\text{Ch}\{B\}} = \text{Ch}\{A_2 | B\}.$$

If

$$\frac{\text{Ch}\{A_1 \cap B\}}{\text{Ch}\{B\}} \leq 0.5 \leq \frac{\text{Ch}\{A_2 \cap B\}}{\text{Ch}\{B\}},$$

then $\text{Ch}\{A_1 | B\} \leq 0.5 \leq \text{Ch}\{A_2 | B\}$. If

$$0.5 < \frac{\text{Ch}\{A_1 \cap B\}}{\text{Ch}\{B\}} \leq \frac{\text{Ch}\{A_2 \cap B\}}{\text{Ch}\{B\}},$$

then we have

$$\text{Ch}\{A_1 | B\} = \left(1 - \frac{\text{Ch}\{A_1^c \cap B\}}{\text{Ch}\{B\}}\right) \vee 0.5 \leq \left(1 - \frac{\text{Ch}\{A_2^c \cap B\}}{\text{Ch}\{B\}}\right) \vee 0.5 = \text{Ch}\{A_2 | B\}.$$

This means that $\text{Ch}\{\cdot | B\}$ is increasing. For any event A, if

$$\frac{\text{Ch}\{A \cap B\}}{\text{Ch}\{B\}} \geq 0.5, \quad \frac{\text{Ch}\{A^c \cap B\}}{\text{Ch}\{B\}} \geq 0.5,$$

then we have $\text{Ch}\{A | B\} + \text{Ch}\{A^c | B\} = 0.5 + 0.5 = 1$ immediately. Otherwise, without loss of generality, suppose

$$\frac{\text{Ch}\{A \cap B\}}{\text{Ch}\{B\}} < 0.5 < \frac{\text{Ch}\{A^c \cap B\}}{\text{Ch}\{B\}},$$

then we have

$$\text{Ch}\{A|B\} + \text{Ch}\{A^c|B\} = \frac{\text{Ch}\{A \cap B\}}{\text{Ch}\{B\}} + \left(1 - \frac{\text{Ch}\{A \cap B\}}{\text{Ch}\{B\}}\right) = 1.$$

That is, $\text{Ch}\{\cdot|B\}$ is self-dual. Finally, for any countable sequence $\{A_i\}$ of events, if $\text{Ch}\{A_i|B\} < 0.5$ for all i, it follows from the countable subadditivity of chance measure that

$$\text{Ch}\left\{\bigcup_{i=1}^{\infty} A_i \cap B\right\} \le \frac{\text{Ch}\left\{\bigcup_{i=1}^{\infty} A_i \cap B\right\}}{\text{Ch}\{B\}} \le \frac{\sum_{i=1}^{\infty} \text{Ch}\{A_i \cap B\}}{\text{Ch}\{B\}} = \sum_{i=1}^{\infty} \text{Ch}\{A_i|B\}.$$

Suppose there is one term greater than 0.5, say

$$\text{Ch}\{A_1|B\} \ge 0.5, \quad \text{Ch}\{A_i|B\} < 0.5, \quad i = 2, 3, \cdots$$

If $\text{Ch}\{\cup_i A_i|B\} = 0.5$, then we immediately have

$$\text{Ch}\left\{\bigcup_{i=1}^{\infty} A_i \cap B\right\} \le \sum_{i=1}^{\infty} \text{Ch}\{A_i|B\}.$$

If $\text{Ch}\{\cup_i A_i|B\} > 0.5$, we may prove the above inequality by the following facts:

$$A_1^c \cap B \subset \bigcup_{i=2}^{\infty}(A_i \cap B) \cup \left(\bigcap_{i=1}^{\infty} A_i^c \cap B\right),$$

$$\text{Ch}\{A_1^c \cap B\} \le \sum_{i=2}^{\infty} \text{Ch}\{A_i \cap B\} + \text{Ch}\left\{\bigcap_{i=1}^{\infty} A_i^c \cap B\right\},$$

$$\text{Ch}\left\{\bigcup_{i=1}^{\infty} A_i|B\right\} = 1 - \frac{\text{Ch}\left\{\bigcap_{i=1}^{\infty} A_i^c \cap B\right\}}{\text{Ch}\{B\}},$$

$$\sum_{i=1}^{\infty} \text{Ch}\{A_i|B\} \ge 1 - \frac{\text{Ch}\{A_1^c \cap B\}}{\text{Ch}\{B\}} + \frac{\sum_{i=2}^{\infty} \text{Ch}\{A_i \cap B\}}{\text{Ch}\{B\}}.$$

If there are at least two terms greater than 0.5, then the countable subadditivity is clearly true. Thus $\text{Ch}\{\cdot|B\}$ is countably subadditive. Hence the theorem is verified.

Definition C.13 *Let ξ be a hybrid variable on $(\Theta, \mathcal{P}, \text{Cr}) \times (\Omega, \mathcal{A}, \text{Pr})$. A conditional hybrid variable of ξ given B is a measurable function $\xi|_B$ from the conditional chance space to the set of real numbers such that*

$$\xi|_B(\theta, \omega) \equiv \xi(\theta, \omega), \quad \forall(\theta, \omega) \in \Theta \times \Omega. \tag{C.39}$$

Definition C.14 *(Li and Liu [101]) The conditional chance distribution Φ: $\Re \to [0,1]$ of a hybrid variable ξ given B is defined by*

$$\Phi(x|B) = \text{Ch}\{\xi \leq x|B\} \qquad (C.40)$$

provided that $\text{Ch}\{B\} > 0$.

Definition C.15 *(Li and Liu [101]) Let ξ be a hybrid variable. Then the conditional expected value of ξ given B is defined by*

$$E[\xi|B] = \int_0^{+\infty} \text{Ch}\{\xi \geq r|B\}\mathrm{d}r - \int_{-\infty}^0 \text{Ch}\{\xi \leq r|B\}\mathrm{d}r \qquad (C.41)$$

provided that at least one of the two integrals is finite.

Appendix D

Logic

Propositions, connective symbols, formulas, truth functions, truth value are basic concepts in logic. This appendix introduces the probabilistic logic, credibilistic logic, and hybrid logic.

D.1   Probabilistic Logic

Probabilistic logic was proposed by Nilsson [171] as a generalization of logic for dealing with random knowledge. A *random proposition* is a statement with probability value belonging to $[0, 1]$. A *random formula* is a finite sequence of random propositions and connective symbols that must make sense. Let X be a random formula. Then the truth value of X is defined as the probability that the random formula X is true, i.e.,

$$T(X) = \Pr\{X = 1\}. \tag{D.1}$$

If X is a random formula containing random propositions $\xi_1, \xi_2, \cdots, \xi_n$ whose truth function is $f(x_1, x_2, \cdots, x_n)$, then the truth value of X is

$$T(X) = \sum_{f(x_1, x_2, \cdots, x_n) = 1} p(x_1, x_2, \cdots, x_n) \tag{D.2}$$

where $p(x_1, x_2, \cdots, x_n)$ is the joint probability mass function of $(\xi_1, \xi_2, \cdots, \xi_n)$.

Example D.1: Let ξ and η be independent random propositions with probability values a and b, respectively. Then we have
(a) $T(\xi) = \Pr\{\xi = 1\} = a$;
(b) $T(\neg \xi) = \Pr\{\xi = 0\} = 1 - a$;
(c) $T(\xi \wedge \eta) = \Pr\{\xi \wedge \eta = 1\} = \Pr\{(\xi = 1) \cap (\eta = 1)\} = a \cdot b$;
(d) $T(\xi \vee \eta) = \Pr\{\xi \vee \eta = 1\} = \Pr\{(\xi = 1) \cup (\eta = 1)\} = a + b - a \cdot b$;
(e) $T(\xi \rightarrow \eta) = \Pr\{(\xi \rightarrow \eta) = 1\} = \Pr\{(\xi = 0) \cup (\eta = 1)\} = 1 - a + a \cdot b$.

Remark D.1: Probabilistic logic is consistent with the law of contradiction and law of excluded middle, i.e.,

$$T(X \wedge \neg X) = 0, \quad T(X \vee \neg X) = 1, \quad T(X) + T(\neg X) = 1. \qquad \text{(D.3)}$$

D.2   Credibilistic Logic

Credibilistic logic was designed by Li and Liu [100] as a generalization of logic for dealing with fuzzy knowledge. A *fuzzy proposition* is a statement with credibility value belonging to $[0, 1]$. A *fuzzy formula* is a finite sequence of fuzzy propositions and connective symbols that must make sense. Let X be a fuzzy formula. Then the truth value of X is defined as the credibility that the fuzzy formula X is true, i.e.,

$$T(X) = \text{Cr}\{X = 1\}. \qquad \text{(D.4)}$$

If X is a fuzzy formula containing fuzzy propositions $\xi_1, \xi_2, \cdots, \xi_n$ whose truth function is $f(x_1, x_2, \cdots, x_n)$, then the truth value of X is

$$T(X) = \frac{1}{2} \left(\max_{f(\boldsymbol{x})=1} \mu(\boldsymbol{x}) + 1 - \max_{f(\boldsymbol{x})=0} \mu(\boldsymbol{x}) \right) \qquad \text{(D.5)}$$

where $\boldsymbol{x} = (x_1, x_2, \cdots, x_n)$, and $\mu(\boldsymbol{x})$ is the joint membership function of $(\xi_1, \xi_2, \cdots, \xi_n)$.

Example D.2: Let ξ and η be (not necessarily independent) fuzzy propositions with credibility values a and b, respectively. Then we have
(a) $T(\xi) = \text{Cr}\{\xi = 1\} = a$;
(b) $T(\neg \xi) = \text{Cr}\{\xi = 0\} = 1 - a$;
(c) $T(\xi \wedge \eta) = \text{Cr}\{(\xi = 1) \cap (\eta = 1)\} = a \wedge b$ provided that $a + b > 1$;
(d) $T(\xi \vee \eta) = \text{Cr}\{(\xi = 1) \cup (\eta = 1)\} = a \vee b$ provided that $a + b < 1$;
(e) $T(\xi \to \eta) = \text{Cr}\{(\xi = 0) \cup (\eta = 1)\} = (1 - a) \vee b$ provided that $a > b$.

Example D.3: Let ξ and η be independent fuzzy propositions with credibility values a and b, respectively. Then we have
(a) $T(\xi \wedge \eta) = \text{Cr}\{(\xi = 1) \cap (\eta = 1)\} = a \wedge b$;
(b) $T(\xi \vee \eta) = \text{Cr}\{(\xi = 1) \cup (\eta = 1)\} = a \vee b$;
(c) $T(\xi \to \eta) = \text{Cr}\{(\xi = 0) \cup (\eta = 1)\} = (1 - a) \vee b$.

Remark D.2: Credibilistic logic is consistent with the law of contradiction and law of excluded middle, i.e.,

$$T(X \wedge \neg X) = 0, \quad T(X \vee \neg X) = 1, \quad T(X) + T(\neg X) = 1. \qquad \text{(D.6)}$$

This is the key point that credibilistic logic is different from possibilistic logic.

D.3   Hybrid Logic

Hybrid logic was designed by Li and Liu [99] in 2009 as a generalization of logic for dealing with fuzzy knowledge and random knowledge simultaneously. A *hybrid formula* is a finite sequence of fuzzy propositions, random propositions and connective symbols that must make sense. Let X be a hybrid formula. Then the truth value of X is defined as the chance that the hybrid formula X is true, i.e.,

$$T(X) = \mathrm{Ch}\{X = 1\}. \tag{D.7}$$

If X is a hybrid formula containing fuzzy propositions $\xi_1, \xi_2, \cdots, \xi_n$ and random propositions $\eta_1, \eta_2, \cdots, \eta_m$ whose truth function is f, then the truth value of X is

$$T(X) = \begin{cases} \sup_{\boldsymbol{x}} \left(\dfrac{\mu(\boldsymbol{x})}{2} \wedge \displaystyle\sum_{f(\boldsymbol{x}, \boldsymbol{y})=1} \phi(\boldsymbol{y}) \right), \\[2em] \qquad \text{if } \sup_{\boldsymbol{x}} \left(\dfrac{\mu(\boldsymbol{x})}{2} \wedge \displaystyle\sum_{f(\boldsymbol{x}, \boldsymbol{y})=1} \phi(\boldsymbol{y}) \right) < 0.5 \\[2em] 1 - \sup_{\boldsymbol{x}} \left(\dfrac{\mu(\boldsymbol{x})}{2} \wedge \displaystyle\sum_{f(\boldsymbol{x}, \boldsymbol{y})=0} \phi(\boldsymbol{y}) \right), \\[2em] \qquad \text{if } \sup_{\boldsymbol{x}} \left(\dfrac{\mu(\boldsymbol{x})}{2} \wedge \displaystyle\sum_{f(\boldsymbol{x}, \boldsymbol{y})=1} \phi(\boldsymbol{y}) \right) \geq 0.5 \end{cases} \tag{D.8}$$

where $\boldsymbol{x} = (x_1, x_2, \cdots, x_n)$, $\boldsymbol{y} = (y_1, y_2, \cdots, y_m)$, $\mu(\boldsymbol{x})$ is the joint membership function of $(\xi_1, \xi_2, \cdots, \xi_n)$, and $\phi(\boldsymbol{y})$ is the joint probability mass function of $(\eta_1, \eta_2, \cdots, \eta_m)$.

Example D.4: Let ξ, κ be independent fuzzy propositions with credibility values a, c, and let η, τ be independent random propositions with probability values b, d, respectively. Then we have
(a) $T(\xi \wedge \eta) = \mathrm{Ch}\{(\xi = 1) \cap (\eta = 1)\} = a \wedge b$;
(b) $T(\xi \vee \eta) = \mathrm{Ch}\{(\xi = 1) \cup (\eta = 1)\} = a \vee b$;
(c) $T(\xi \to \tau) = \mathrm{Ch}\{(\xi = 0) \cup (\tau = 1)\} = (1 - a) \vee d$;
(d) $T(\eta \to \kappa) = \mathrm{Ch}\{(\eta = 0) \cup (\kappa = 1)\} = (1 - b) \vee c$.
Furthermore, we have

$$\begin{aligned} T(\xi \wedge \eta \to \kappa) &= \mathrm{Ch}\{(\xi \wedge \eta = 0) \cup (\kappa = 1)\} \\ &= 1 - \mathrm{Ch}\{(\xi = 1) \cap (\eta = 1) \cap (\kappa = 0)\} \\ &= 1 - \mathrm{Cr}\{(\xi = 1) \cap (\kappa = 0)\} \wedge \mathrm{Pr}\{\eta = 1\} \\ &= 1 - a \wedge (1 - c) \wedge b; \end{aligned}$$

$$
\begin{aligned}
T(\xi \wedge \eta \to \tau) &= \mathrm{Ch}\{(\xi \wedge \eta = 0) \cup (\tau = 1)\} \\
&= 1 - \mathrm{Ch}\{(\xi = 1) \cap (\eta = 1) \cap (\tau = 0)\} \\
&= 1 - \mathrm{Cr}\{\xi = 1\} \wedge \mathrm{Pr}\{(\eta = 1) \cap (\tau = 0)\} \\
&= 1 - a \wedge (b - bd).
\end{aligned}
$$

Remark D.3: Hybrid logic is consistent with the law of contradiction and law of excluded middle, i.e.,

$$
T(X \wedge \neg X) = 0, \quad T(X \vee \neg X) = 1, \quad T(X) + T(\neg X) = 1. \qquad \text{(D.9)}
$$

Appendix E

Inference

Inference is a process of deriving logical conclusions from known knowledge or evidence. This appendix introduces random inference, fuzzy inference and hybrid inference.

E.1   Random Inference

Random inference is a process of deriving logical conclusions from random knowledge or evidence via the tool of conditional probability.

Definition 1 *A random set is a measurable function ξ from a probability space $(\Omega, \mathcal{A}, \mathrm{Pr})$ to a collection of sets of real numbers, i.e., for any Borel set B, the set*

$$\{\omega \in \Omega \mid \xi(\omega) \cap B \neq \emptyset\} \tag{E.1}$$

is an event.

Definition 2 *Let ξ and η be random sets on the probability space $(\Omega, \mathcal{A}, \mathrm{Pr})$. Then the complement ξ^c of the random set ξ is*

$$\xi^c(\omega) = (\xi(\omega))^c, \quad \forall \omega \in \Omega. \tag{E.2}$$

The union $\xi \cup \eta$ of random sets ξ and η is

$$(\xi \cup \eta)(\omega) = \xi(\omega) \cup \eta(\omega), \quad \forall \omega \in \Omega. \tag{E.3}$$

The intersection $\xi \cap \eta$ of random sets ξ and η is

$$(\xi \cap \eta)(\omega) = \xi(\omega) \cap \eta(\omega), \quad \forall \omega \in \Omega. \tag{E.4}$$

It is clear that

$$\xi \cup \xi^c \equiv \Re, \quad \xi \cap \xi^c \equiv \emptyset \tag{E.5}$$

for any radom set ξ.

Definition 3 *Let ξ be a random set on $(\Omega, \mathcal{A}, \mathrm{Pr})$. A conditional random set of ξ given B is a measurable function $\xi|_B$ from the conditional probability space to a collection of sets of real numbers such that*

$$\xi|_B(\omega) \equiv \xi(\omega), \quad \forall \omega \in \Omega \tag{E.6}$$

in the sense of classical set theory.

Definition 4 *Let ξ and ξ^* be two random sets. Then the event $\xi^* \rhd \xi$ (i.e., ξ^* matches ξ) is defined as*

$$\{\xi^* \rhd \xi\} = \begin{cases} \{\xi^* \subset \xi\}, & \text{if } \mathrm{Pr}\{\xi^* \subset \xi\} > \mathrm{Pr}\{\xi^* \subset \xi^c\} \\ \{\xi^* \not\subset \xi^c\}, & \text{if } \mathrm{Pr}\{\xi^* \subset \xi\} \leq \mathrm{Pr}\{\xi^* \subset \xi^c\}. \end{cases} \tag{E.7}$$

Definition 5 *Let ξ and ξ^* be two random sets. Then the matching degree of $\xi^* \rhd \xi$ is defined as*

$$\mathrm{Pr}\{\xi^* \rhd \xi\} = \begin{cases} \mathrm{Pr}\{\xi^* \subset \xi\}, & \text{if } \mathrm{Pr}\{\xi^* \subset \xi\} > \mathrm{Pr}\{\xi^* \subset \xi^c\} \\ \mathrm{Pr}\{\xi^* \not\subset \xi^c\}, & \text{if } \mathrm{Pr}\{\xi^* \subset \xi\} \leq \mathrm{Pr}\{\xi^* \subset \xi^c\}. \end{cases} \tag{E.8}$$

Inference Rule E.1 *Let $\mathbb{X}$ and $\mathbb{Y}$ be two concepts. Assume a rule "if $\mathbb{X}$ is a random set ξ then $\mathbb{Y}$ is a random set η". From $\mathbb{X}$ is a random set ξ^* we infer that $\mathbb{Y}$ is a random set*

$$\eta^* = \eta|_{\xi^* \rhd \xi} \tag{E.9}$$

which is the conditional random set of η given $\xi^* \rhd \xi$. The inference rule is represented by

$$\begin{array}{c} \text{Rule: If } \mathbb{X} \text{ is } \xi \text{ then } \mathbb{Y} \text{ is } \eta \\ \text{From: } \mathbb{X} \text{ is } \xi^* \\ \hline \text{Infer: } \mathbb{Y} \text{ is } \eta^* = \eta|_{\xi^* \rhd \xi} \end{array} \tag{E.10}$$

E.2   Fuzzy Inference

Fuzzy inference was proposed by Liu [130] in 2009 as a process of deriving logical conclusions from fuzzy knowledge or evidence via the tool of conditional credibility.

Before introducing fuzzy inference, let us recall the concept of fuzzy set. An ordinary set is defined as a collection of elements. Each single element can either belong or not belong to the set. However, in many cases, the membership is not clear, for example, "young man", "many people", "high mountain", "great river", "large number", "about 100 meters". They are not tractable by the classical set theory. In order to deal with them, Zadeh [229] introduced the concept of fuzzy set via membership function in 1965: *A fuzzy set is defined by its membership function μ which assigns to each element x a real number $\mu(x)$ in the interval $[0,1]$, where the value of $\mu(x)$ represents*

the grade of membership of x in the fuzzy set. Thus, the nearer the value of $\mu(x)$ is unity, the higher the grade of membership of x in the fuzzy set.

As a revision of Zadeh's fuzzy set, Liu [127] redefined a fuzzy set as a function from a credibility space to a collection of sets in 2007. This new version of fuzzy set is just a counterpart of random set that is a measurable function from a probability space to a collection of sets.

Definition 6 *(Liu [127]) A fuzzy set is a function ξ from a credibility space $(\Theta, \mathcal{P}, \mathrm{Cr})$ to a collection of sets of real numbers.*

Definition 7 *A fuzzy set ξ on $(\Theta, \mathcal{P}, \mathrm{Cr})$ is said to have a membership function $\mu(x)$ with $0 \leq \mu(x) \leq 1$ if*

$$\xi(\theta) = \{x \in \Re \mid \mu(x) \geq (2\mathrm{Cr}\{\theta\}) \wedge 1\} \tag{E.11}$$

for each $\theta \in \Theta$.

Note that Zadeh's fuzzy set is a special type of fuzzy set in the sense of Definition 6. However, the latter does not necessarily have a membership function.

Example E.1: Let $(\Theta, \mathcal{P}, \mathrm{Cr})$ be a credibility space with $\Theta = \{\theta_1, \theta_2, \theta_3\}$, $\mathrm{Cr}\{\theta_1\} = 0.6$, $\mathrm{Cr}\{\theta_2\} = 0.4$ and $\mathrm{Cr}\{\theta_3\} = 0.2$. Then

$$\xi(\theta) = \begin{cases} [1,3], & \text{if } \theta = \theta_1 \\ [2,4], & \text{if } \theta = \theta_2 \\ [3,5], & \text{if } \theta = \theta_3 \end{cases}$$

is a fuzzy set whose membership function does not exist.

Definition 8 *Let ξ and η be fuzzy sets on the credibility space $(\Theta, \mathcal{P}, \mathrm{Cr})$. Then the complement ξ^c of the fuzzy set ξ is*

$$\xi^c(\theta) = (\xi(\theta))^c, \quad \forall \theta \in \Theta. \tag{E.12}$$

The union $\xi \cup \eta$ of fuzzy sets ξ and η is

$$(\xi \cup \eta)(\theta) = \xi(\theta) \cup \eta(\theta), \quad \forall \theta \in \Theta. \tag{E.13}$$

The intersection $\xi \cap \eta$ of fuzzy sets ξ and η is

$$(\xi \cap \eta)(\theta) = \xi(\theta) \cap \eta(\theta), \quad \forall \theta \in \Theta. \tag{E.14}$$

It is clear that

$$\xi \cup \xi^c \equiv \Re, \quad \xi \cap \xi^c \equiv \emptyset \tag{E.15}$$

for any fuzzy set ξ. This point is very different from Zadeh's fuzzy set.

Definition 9 *Let ξ be a fuzzy set on $(\Theta, \mathcal{P}, \mathrm{Cr})$. A conditional fuzzy set of ξ given B is a function $\xi|_B$ from the conditional credibility space $(\Theta, \mathcal{P}, \mathrm{Cr}\{\cdot|B\})$ to a collection of sets of real numbers such that*

$$\xi|_B(\theta) \equiv \xi(\theta), \quad \forall \theta \in \Theta \tag{E.16}$$

in the sense of classical set theory.

Definition 10 *Let ξ and ξ^* be two fuzzy sets. Then the event $\xi^* \rhd \xi$ (i.e., ξ^* matches ξ) is defined as*

$$\{\xi^* \rhd \xi\} = \begin{cases} \{\xi^* \subset \xi\}, & \textit{if } \mathrm{Cr}\{\xi^* \subset \xi\} > \mathrm{Cr}\{\xi^* \subset \xi^c\} \\ \{\xi^* \not\subset \xi^c\}, & \textit{if } \mathrm{Cr}\{\xi^* \subset \xi\} \leq \mathrm{Cr}\{\xi^* \subset \xi^c\}. \end{cases} \tag{E.17}$$

That is, we take the event of $\{\xi^* \subset \xi\}$ or $\{\xi^* \not\subset \xi^c\}$ such that its credibility measure is as closed to 0.5 as possible. This idea coincides with the maximum uncertainty principle.

Definition 11 *Let ξ and ξ^* be two fuzzy sets. Then the matching degree of $\xi^* \rhd \xi$ is defined as*

$$\mathrm{Cr}\{\xi^* \rhd \xi\} = \begin{cases} \mathrm{Cr}\{\xi^* \subset \xi\}, & \textit{if } \mathrm{Cr}\{\xi^* \subset \xi\} > \mathrm{Cr}\{\xi^* \subset \xi^c\} \\ \mathrm{Cr}\{\xi^* \not\subset \xi^c\}, & \textit{if } \mathrm{Cr}\{\xi^* \subset \xi\} \leq \mathrm{Cr}\{\xi^* \subset \xi^c\}. \end{cases} \tag{E.18}$$

Fuzzy inference has been studied by many scholars and a lot of inference rules have been suggested. For example, Zadeh's compositional rule of inference, Lukasiewicz's inference rule, and Mamdani's inference rule are widely used in fuzzy inference systems. However, instead of using the above inference rules, we will introduce an inference rule proposed by Liu in 2009.

Inference Rule E.2 *(Liu [130]) Let $\mathbb{X}$ and $\mathbb{Y}$ be two concepts. Assume a rule "if $\mathbb{X}$ is a fuzzy set ξ then $\mathbb{Y}$ is a fuzzy set η". From $\mathbb{X}$ is a fuzzy set ξ^* we infer that $\mathbb{Y}$ is a fuzzy set*

$$\eta^* = \eta|_{\xi^* \rhd \xi} \tag{E.19}$$

which is the conditional fuzzy set of η given $\xi^* \rhd \xi$. The inference rule is represented by

$$\begin{array}{l} \text{Rule: If } \mathbb{X} \text{ is } \xi \text{ then } \mathbb{Y} \text{ is } \eta \\ \text{From: } \mathbb{X} \text{ is } \xi^* \\ \hline \text{Infer: } \mathbb{Y} \text{ is } \eta^* = \eta|_{\xi^* \rhd \xi} \end{array} \tag{E.20}$$

Example E.2: Let ξ, ξ^* and η be independent fuzzy sets with membership functions μ, μ^* and ν, respectively. Then inference rule E.2 yields that η^*

has a membership function

$$
\nu^*(y) = \begin{cases} \dfrac{2\nu(y)}{\sup\limits_{x\in\Re} \mu(x) \wedge \mu^*(x)} \wedge 1, & \text{if } \sup\limits_{x\in\Re} \mu(x) \wedge \mu^*(x) < 1 \\[3ex] \dfrac{2\nu(y)}{2 - \sup\limits_{x\in\Re}\{\mu(x)|\mu(x) < \mu^*(x)\}} \wedge 1, & \text{if } \sup\limits_{x\in\Re} \mu(x) \wedge \mu^*(x) = 1. \end{cases}
$$

Example E.3: Let ξ and η be independent fuzzy sets with membership functions μ and ν, respectively. If ξ^* is a constant a, then inference rule E.2 yields that η^* has a membership function

$$
\nu^*(y) = \begin{cases} \dfrac{2\nu(y)}{\mu(a)} \wedge 1, & \text{if } \mu(a) < 1 \\[2ex] \nu(y), & \text{if } \mu(a) = 1. \end{cases}
$$

For more expressions of inference rule by membership functions, the interested reader may consult Gao [43].

E.3   Hybrid Inference

Hybrid inference was proposed by Liu [130] in 2009 as a process of deriving logical conclusions from hybrid knowledge or evidence via the tool of conditional chance.

Definition 12 *(Liu [127]) A hybrid set is a measurable function ξ from a chance space $(\Theta, \mathcal{P}, \mathrm{Cr}) \times (\Omega, \mathcal{A}, \mathrm{Pr})$ to a collection of sets of real numbers, i.e., for any Borel set B, the set*

$$
\{(\theta, \omega) \in \Theta \times \Omega \mid \xi(\theta, \omega) \cap B \neq \emptyset\} \tag{E.21}
$$

is an event.

Definition 13 *Let ξ and η be hybrid sets on the chance space $(\Theta, \mathcal{P}, \mathrm{Cr}) \times (\Omega, \mathcal{A}, \mathrm{Pr})$. Then the complement ξ^c of the hybrid set ξ is*

$$
\xi^c(\theta, \omega) = (\xi(\theta, \omega))^c, \quad \forall (\theta, \omega) \in \Theta \times \Omega. \tag{E.22}
$$

The union $\xi \cup \eta$ of hybrid sets ξ and η is

$$
(\xi \cup \eta)(\theta, \omega) = \xi(\theta, \omega) \cup \eta(\theta, \omega), \quad \forall (\theta, \omega) \in \Theta \times \Omega. \tag{E.23}
$$

The intersection $\xi \cap \eta$ of hybrid sets ξ and η is

$$
(\xi \cap \eta)(\theta, \omega) = \xi(\theta, \omega) \cap \eta(\theta, \omega), \quad \forall (\theta, \omega) \in \Theta \times \Omega. \tag{E.24}
$$

It is clear that

$$\xi \cup \xi^c \equiv \Re, \quad \xi \cap \xi^c \equiv \emptyset \tag{E.25}$$

for any hybrid set ξ.

Definition 14 *Let ξ be a hybrid set on $(\Theta, \mathcal{P}, \mathrm{Cr}) \times (\Omega, \mathcal{A}, \mathrm{Pr})$. A conditional hybrid set of ξ given B is a measurable function $\xi|_B$ from the conditional chance space to a collection of sets of real numbers such that*

$$\xi|_B(\theta, \omega) \equiv \xi(\theta, \omega), \quad \forall (\theta, \omega) \in \Theta \times \Omega \tag{E.26}$$

in the sense of classical set theory.

Definition 15 *Let ξ and ξ^* be two hybrid sets. Then the event $\xi^* \triangleright \xi$ (i.e., ξ^* matches ξ) is defined as*

$$\{\xi^* \triangleright \xi\} = \begin{cases} \{\xi^* \subset \xi\}, & \text{if } \mathrm{Ch}\{\xi^* \subset \xi\} > \mathrm{Ch}\{\xi^* \subset \xi^c\} \\ \{\xi^* \not\subset \xi^c\}, & \text{if } \mathrm{Ch}\{\xi^* \subset \xi\} \le \mathrm{Ch}\{\xi^* \subset \xi^c\}. \end{cases} \tag{E.27}$$

Definition 16 *Let ξ and ξ^* be two hybrid sets. Then the matching degree of $\xi^* \triangleright \xi$ is defined as*

$$\mathrm{Ch}\{\xi^* \triangleright \xi\} = \begin{cases} \mathrm{Ch}\{\xi^* \subset \xi\}, & \text{if } \mathrm{Ch}\{\xi^* \subset \xi\} > \mathrm{Ch}\{\xi^* \subset \xi^c\} \\ \mathrm{Ch}\{\xi^* \not\subset \xi^c\}, & \text{if } \mathrm{Ch}\{\xi^* \subset \xi\} \le \mathrm{Ch}\{\xi^* \subset \xi^c\}. \end{cases} \tag{E.28}$$

Inference Rule E.3 *(Liu [130]) Let $\mathbb{X}$ and $\mathbb{Y}$ be two concepts. Assume a rule "if $\mathbb{X}$ is a hybrid set ξ then $\mathbb{Y}$ is a hybrid set η". From $\mathbb{X}$ is a hybrid set ξ^* we infer that $\mathbb{Y}$ is a hybrid set*

$$\eta^* = \eta|_{\xi^* \triangleright \xi} \tag{E.29}$$

which is the conditional hybrid set of η given $\xi^* \triangleright \xi$. The inference rule is represented by

$$\frac{\begin{array}{l} \text{Rule: If } \mathbb{X} \text{ is } \xi \text{ then } \mathbb{Y} \text{ is } \eta \\ \text{From: } \mathbb{X} \text{ is } \xi^* \end{array}}{\text{Infer: } \mathbb{Y} \text{ is } \eta^* = \eta|_{\xi^* \triangleright \xi}} \tag{E.30}$$

Bibliography

[1] Alefeld G, Herzberger J, *Introduction to Interval Computations*, Academic Press, New York, 1983.

[2] Atanassov KT, *Intuitionistic Fuzzy Sets: Theory and Applications*, Physica-Verlag, Heidelberg, 1999.

[3] Bamber D, Goodman IR, Nguyen HT, Extension of the concept of propositional deduction from classical logic to probability: An overview of probability-selection approaches, *Information Sciences*, Vol.131, 195-250, 2001.

[4] Bandemer H, and Nather W, *Fuzzy Data Analysis*, Kluwer, Dordrecht, 1992.

[5] Bedford T, and Cooke MR, *Probabilistic Risk Analysis*, Cambridge University Press, 2001.

[6] Bellman RE, and Zadeh LA, Decision making in a fuzzy environment, *Management Science*, Vol.17, 141-164, 1970.

[7] Bhandari D, and Pal NR, Some new information measures of fuzzy sets, *Information Sciences*, Vol.67, 209-228, 1993.

[8] Black F, and Scholes M, The pricing of option and corporate liabilities, *Journal of Political Economy*, Vol.81, 637-654, 1973.

[9] Bouchon-Meunier B, Mesiar R, Ralescu DA, Linear non-additive set-functions, *International Journal of General Systems*, Vol.33, No.1, 89-98, 2004.

[10] Buckley JJ, Possibility and necessity in optimization, *Fuzzy Sets and Systems*, Vol.25, 1-13, 1988.

[11] Buckley JJ, Stochastic versus possibilistic programming, *Fuzzy Sets and Systems*, Vol.34, 173-177, 1990.

[12] Cadenas JM, and Verdegay JL, Using fuzzy numbers in linear programming, *IEEE Transactions on Systems, Man and Cybernetics–Part B*, Vol.27, No.6, 1016-1022, 1997.

[13] Campos L, and González, A, A subjective approach for ranking fuzzy numbers, *Fuzzy Sets and Systems*, Vol.29, 145-153, 1989.

[14] Campos L, and Verdegay JL, Linear programming problems and ranking of fuzzy numbers, *Fuzzy Sets and Systems*, Vol.32, 1-11, 1989.

[15] Campos FA, Villar J, and Jimenez M, Robust solutions using fuzzy chance constraints, *Engineering Optimization*, Vol.38, No.6, 627-645, 2006.

[16] Chen SJ, and Hwang CL, *Fuzzy Multiple Attribute Decision Making: Methods and Applications*, Springer-Verlag, Berlin, 1992.

[17] Chen XW, and Liu B, Existence and uniqueness theorem for uncertain differential equations, *Fuzzy Optimization and Decision Making*, to be published.

[18] Chen XW, and Ralescu DA, A note on truth value in uncertain logic, *Proceedings of the Eighth International Conference on Information and Management Sciences*, Kunming, China, July 20-28, 2009, pp.739-741. (Also available at http://orsc.edu.cn/online/090211.pdf)

[19] Chen XW, and Dai W, Maximum entropy principle for uncertain variables, http://orsc.edu.cn/online/090618.pdf.

[20] Chen XW, Stability analysis of linear uncertain differential equations, http://orsc.edu.cn/online/091001.pdf.

[21] Chen Y, Fung RYK, Yang J, Fuzzy expected value modelling approach for determining target values of engineering characteristics in QFD, *International Journal of Production Research*, Vol.43, No.17, 3583-3604, 2005.

[22] Chen Y, Fung RYK, Tang JF, Rating technical attributes in fuzzy QFD by integrating fuzzy weighted average method and fuzzy expected value operator, *European Journal of Operational Research*, Vol.174, No.3, 1553-1566, 2006.

[23] Choquet G, Theory of capacities, *Annals de l'Institute Fourier*, Vol.5, 131-295, 1954.

[24] Dai W, and Chen XW, Entropy of function of uncertain variables, http://orsc.edu.cn/online/090805.pdf.

[25] Das B, Maity K, Maiti A, A two warehouse supply-chain model under possibility/necessity/credibility measures, *Mathematical and Computer Modelling*, Vol.46, No.3-4, 398-409, 2007.

[26] De Cooman G, Possibility theory I-III, *International Journal of General Systems*, Vol.25, 291-371, 1997.

[27] De Luca A, and Termini S, A definition of nonprobabilistic entropy in the setting of fuzzy sets theory, *Information and Control*, Vol.20, 301-312, 1972.

[28] Dempster AP, Upper and lower probabilities induced by a multivalued mapping, *Ann. Math. Stat.*, Vol.38, No.2, 325-339, 1967.

[29] Dubois D, and Prade H, Fuzzy logics and generalized modus ponens revisited, *Cybernetics and Systems*, Vol.15, 293-331, 1984.

[30] Dubois D, and Prade H, The mean value of a fuzzy number, *Fuzzy Sets and Systems*, Vol.24, 279-300, 1987.

[31] Elkan C, The paradoxical success of fuzzy logic, *IEEE Expert*, Vol.9, No.4, 3-8, 1994.

[32] Elkan C, The paradoxical controversy over fuzzy logic, *IEEE Expert*, Vol.9, No.4, 47-49, 1994.

[33] Esogbue AO, and Liu B, Reservoir operations optimization via fuzzy criterion decision processes, *Fuzzy Optimization and Decision Making*, Vol.5, No.3, 289-305, 2006.

[34] Feng X, and Liu YK, Measurability criteria for fuzzy random vectors, *Fuzzy Optimization and Decision Making*, Vol.5, No.3, 245-253, 2006.

[35] Feng Y, and Yang LX, A two-objective fuzzy k-cardinality assignment problem, *Journal of Computational and Applied Mathematics*, Vol.197, No.1, 233-244, 2006.

[36] Fung RYK, Chen YZ, and Chen L, A fuzzy expected value-based goal programing model for product planning using quality function deployment, *Engineering Optimization*, Vol.37, No.6, 633-647, 2005.

[37] Gao J, and Liu B, New primitive chance measures of fuzzy random event, *International Journal of Fuzzy Systems*, Vol.3, No.4, 527-531, 2001.

[38] Gao J, and Liu B, Fuzzy multilevel programming with a hybrid intelligent algorithm, *Computers & Mathematics with Applications*, Vol.49, 1539-1548, 2005.

[39] Gao J, and Lu M, Fuzzy quadratic minimum spanning tree problem, *Applied Mathematics and Computation*, Vol.164, No.3, 773-788, 2005.

[40] Gao J, Credibilistic game with fuzzy information, *Journal of Uncertain Systems*, Vol.1, No.1, 74-80, 2007.

[41] Gao J, and Gao X, A new stock model for credibilistic option pricing, *Journal of Uncertain Systems*, Vol.2, No.4, 243-247, 2008.

[42] Gao X, Some properties of continuous uncertain measure, *International Journal of Uncertainty, Fuzziness & Knowledge-Based Systems*, Vol.17, No.3, 419-426, 2009.

[43] Gao X, Some expressions of Liu's inference rule by membership functions, http://orsc.edu.cn/online/090118.pdf.

[44] Gao X, and Ralescu DA, On Liu's inference rule for uncertain systems, *Proceedings of the Eighth International Conference on Information and Management Sciences*, Kunming, China, July 20-28, 2009, pp.958-964. (Also available at http://orsc.edu.cn/online/090121.pdf)

[45] Gao Y, Analysis of k-out-of-n system with uncertain lifetimes, *Proceedings of the Eighth International Conference on Information and Management Sciences*, Kunming, China, July 20-28, 2009, pp.794-797.

[46] Gao Y, Interpolation method for Liu's inference rule with uncertain knowledge, http://orsc.edu.cn/online/090921.pdf.

[47] González, A, A study of the ranking function approach through mean values, *Fuzzy Sets and Systems*, Vol.35, 29-41, 1990.

[48] Guan J, and Bell DA, *Evidence Theory and its Applications*, North-Holland, Amsterdam, 1991.

[49] Guo R, Zhao R, Guo D, and Dunne T, Random fuzzy variable modeling on repairable system, *Journal of Uncertain Systems*, Vol.1, No.3, 222-234, 2007.

[50] Ha MH, Wang XZ, and Yang LZ, Sequences of (S) fuzzy integrable functions, *Fuzzy Sets and Systems*, Vol.138, No.3, 507-522, 2003.

[51] Ha MH, Li Y, and Wang XF, Fuzzy knowledge representation and reasoning using a generalized fuzzy petri net and a similarity measure, *Soft Computing*, Vol.11, No.4, 323-327, 2007.

[52] Hansen E, *Global Optimization Using Interval Analysis*, Marcel Dekker, New York, 1992.

[53] He Y, and Xu J, A class of random fuzzy programming model and its application to vehicle routing problem, *World Journal of Modelling and Simulation*, Vol.1, No.1, 3-11, 2005.

[54] Heilpern S, The expected value of a fuzzy number, *Fuzzy Sets and Systems*, Vol.47, 81-86, 1992.

[55] Higashi M, and Klir GJ, On measures of fuzziness and fuzzy complements, *International Journal of General Systems*, Vol.8, 169-180, 1982.

[56] Higashi M, and Klir GJ, Measures of uncertainty and information based on possibility distributions, *International Journal of General Systems*, Vol.9, 43-58, 1983.

[57] Hisdal E, Conditional possibilities independence and noninteraction, *Fuzzy Sets and Systems*, Vol.1, 283-297, 1978.

[58] Hisdal E, *Logical Structures for Representation of Knowledge and Uncertainty*, Physica-Verlag, Heidelberg, 1998.

[59] Hong DH, Renewal process with T-related fuzzy inter-arrival times and fuzzy rewards, *Information Sciences*, Vol.176, No.16, 2386-2395, 2006.

[60] Hu BG, Mann GKI, and Gosine RG, New methodology for analytical and optimal design of fuzzy PID controllers, *IEEE Transactions on Fuzzy Systems*, Vol.7, No.5, 521-539, 1999.

[61] Hu LJ, Wu R, and Shao S, Analysis of dynamical systems whose inputs are fuzzy stochastic processes, *Fuzzy Sets and Systems*, Vol.129, No.1, 111-118, 2002.

[62] Inuiguchi M, and Ramík J, Possibilistic linear programming: A brief review of fuzzy mathematical programming and a comparison with stochastic programming in portfolio selection problem, *Fuzzy Sets and Systems*, Vol.111, No.1, 3-28, 2000.

[63] Ishibuchi H, and Tanaka H, Multiobjective programming in optimization of the interval objective function, *European Journal of Operational Research*, Vol.48, 219-225, 1990.

[64] Jaynes ET, Information theory and statistical mechanics, *Physical Reviews*, Vol.106, No.4, 620-630, 1957.

[65] Ji XY, and Shao Z, Model and algorithm for bilevel Newsboy problem with fuzzy demands and discounts, *Applied Mathematics and Computation*, Vol.172, No.1, 163-174, 2006.

[66] Ji XY, and Iwamura K, New models for shortest path problem with fuzzy arc lengths, *Applied Mathematical Modelling*, Vol.31, 259-269, 2007.

[67] John RI, Type 2 fuzzy sets: An appraisal of theory and applications, *International Journal of Uncertainty, Fuzziness & Knowledge-Based Systems*, Vol.6, No.6, 563-576, 1998.

[68] Kacprzyk J, and Esogbue AO, Fuzzy dynamic programming: Main developments and applications, *Fuzzy Sets and Systems*, Vol.81, 31-45, 1996.

[69] Kacprzyk J, *Multistage Fuzzy Control*, Wiley, Chichester, 1997.

[70] Karnik NN, Mendel JM, and Liang Q, Centroid of a type-2 fuzzy set, *Information Sciences*, Vol.132, 195-220, 2001.

[71] Kaufmann A, *Introduction to the Theory of Fuzzy Subsets*, Vol.I, Academic Press, New York, 1975.

[72] Kaufmann A, and Gupta MM, *Fuzzy Mathematical Models in Engineering and Management Science*, 2nd ed, North-Holland, Amsterdam, 1991.

[73] Ke H, and Liu B, Project scheduling problem with stochastic activity duration times, *Applied Mathematics and Computation*, Vol.168, No.1, 342-353, 2005.

[74] Ke H, and Liu B, Project scheduling problem with mixed uncertainty of randomness and fuzziness, *European Journal of Operational Research*, Vol.183, No.1, 135-147, 2007.

[75] Ke H, and Liu B, Fuzzy project scheduling problem and its hybrid intelligent algorithm, *Applied Mathematical Modelling*, Vol.34, No.2, 301-308, 2010.

[76] Klement EP, Puri ML, and Ralescu DA, Limit theorems for fuzzy random variables, *Proceedings of the Royal Society of London Series A*, Vol.407, 171-182, 1986.

[77] Klir GJ, and Folger TA, *Fuzzy Sets, Uncertainty, and Information*, Prentice-Hall, Englewood Cliffs, NJ, 1980.

[78] Klir GJ, and Yuan B, *Fuzzy Sets and Fuzzy Logic: Theory and Applications*, Prentice-Hall, New Jersey, 1995.

[79] Knopfmacher J, On measures of fuzziness, *Journal of Mathematical Analysis and Applications*, Vol.49, 529-534, 1975.

[80] Kosko B, Fuzzy entropy and conditioning, *Information Sciences*, Vol.40, 165-174, 1986.

[81] Kruse R, and Meyer KD, *Statistics with Vague Data*, D. Reidel Publishing Company, Dordrecht, 1987.

[82] Kwakernaak H, Fuzzy random variables–I: Definitions and theorems, *Information Sciences*, Vol.15, 1-29, 1978.

[83] Kwakernaak H, Fuzzy random variables–II: Algorithms and examples for the discrete case, *Information Sciences*, Vol.17, 253-278, 1979.

[84] Lai YJ, and Hwang CL, *Fuzzy Multiple Objective Decision Making: Methods and Applications*, Springer-Verlag, New York, 1994.

[85] Lee ES, Fuzzy multiple level programming, *Applied Mathematics and Computation*, Vol.120, 79-90, 2001.

[86] Lee KH, *First Course on Fuzzy Theory and Applications*, Springer-Verlag, Berlin, 2005.

[87] Lertworasirkul S, Fang SC, Joines JA, and Nuttle HLW, Fuzzy data envelopment analysis (DEA): a possibility approach, *Fuzzy Sets and Systems*, Vol.139, No.2, 379-394, 2003.

[88] Li DY, Cheung DW, and Shi XM, Uncertainty reasoning based on cloud models in controllers, *Computers and Mathematics with Appications*, Vol.35, No.3, 99-123, 1998.

[89] Li DY, and Du Y, *Artificial Intelligence with Uncertainty*, National Defense Industry Press, Beijing, 2005.

[90] Li J, Xu J, and Gen M, A class of multiobjective linear programming model with fuzzy random coefficients, *Mathematical and Computer Modelling*, Vol.44, Nos.11-12, 1097-1113, 2006.

[91] Li P, and Liu B, Entropy of credibility distributions for fuzzy variables, *IEEE Transactions on Fuzzy Systems*, Vol.16, No.1, 123-129, 2008.

[92] Li SM, Ogura Y, and Kreinovich V, *Limit Theorems and Applications of Set-Valued and Fuzzy Set-Valued Random Variables*, Kluwer, Boston, 2002.

[93] Li X, and Liu B, The independence of fuzzy variables with applications, *International Journal of Natural Sciences & Technology*, Vol.1, No.1, 95-100, 2006.

[94] Li X, and Liu B, A sufficient and necessary condition for credibility measures, *International Journal of Uncertainty, Fuzziness & Knowledge-Based Systems*, Vol.14, No.5, 527-535, 2006.

[95] Li X, and Liu B, Maximum entropy principle for fuzzy variables, *International Journal of Uncertainty, Fuzziness & Knowledge-Based Systems*, Vol.15, Supp.2, 43-52, 2007.

[96] Li X, and Qin ZF, Expected value and variance of geometric Liu process, *Far East Journal of Experimental and Theoretical Artificial Intelligence*, Vol.1, No.2, 127-135, 2008.

[97] Li X, and Liu B, On distance between fuzzy variables, *Journal of Intelligent & Fuzzy Systems*, Vol.19, No.3, 197-204, 2008.

[98] Li X, and Liu B, Chance measure for hybrid events with fuzziness and randomness, *Soft Computing*, Vol.13, No.2, 105-115, 2009.

[99] Li X, and Liu B, Hybrid logic and uncertain logic, *Journal of Uncertain Systems*, Vol.3, No.2, 83-94, 2009.

[100] Li X, and Liu B, Foundation of credibilistic logic, *Fuzzy Optimization and Decision Making*, Vol.8, No.1, 91-102, 2009.

[101] Li X, and Liu B, Conditional chance measure for hybrid events, Technical Report, 2009.

[102] Li X, and Liu B, Cross-entropy and generalized entropy for fuzzy variables, Technical Report, 2009.

[103] Li YM, and Li SJ, A fuzzy sets theoretic approach to approximate spatial reasoning, *IEEE Transactions on Fuzzy Systems*, Vol.12, No.6, 745-754, 2004.

[104] Liu B, Dependent-chance goal programming and its genetic algorithm based approach, *Mathematical and Computer Modelling*, Vol.24, No.7, 43-52, 1996.

[105] Liu B, and Esogbue AO, Fuzzy criterion set and fuzzy criterion dynamic programming, *Journal of Mathematical Analysis and Applications*, Vol.199, No.1, 293-311, 1996.

[106] Liu B, Dependent-chance programming: A class of stochastic optimization, *Computers & Mathematics with Applications*, Vol.34, No.12, 89-104, 1997.

[107] Liu B, and Iwamura K, Modelling stochastic decision systems using dependent-chance programming, *European Journal of Operational Research*, Vol.101, No.1, 193-203, 1997.

[108] Liu B, and Iwamura K, Chance constrained programming with fuzzy parameters, *Fuzzy Sets and Systems*, Vol.94, No.2, 227-237, 1998.

[109] Liu B, and Iwamura K, A note on chance constrained programming with fuzzy coefficients, *Fuzzy Sets and Systems*, Vol.100, Nos.1-3, 229-233, 1998.

[110] Liu B, Minimax chance constrained programming models for fuzzy decision systems, *Information Sciences*, Vol.112, Nos.1-4, 25-38, 1998.

[111] Liu B, Dependent-chance programming with fuzzy decisions, *IEEE Transactions on Fuzzy Systems*, Vol.7, No.3, 354-360, 1999.

[112] Liu B, and Esogbue AO, *Decision Criteria and Optimal Inventory Processes*, Kluwer, Boston, 1999.

[113] Liu B, *Uncertain Programming*, Wiley, New York, 1999.

[114] Liu B, Dependent-chance programming in fuzzy environments, *Fuzzy Sets and Systems*, Vol.109, No.1, 97-106, 2000.

[115] Liu B, Uncertain programming: A unifying optimization theory in various uncertain environments, *Applied Mathematics and Computation*, Vol.120, Nos.1-3, 227-234, 2001.

[116] Liu B, and Iwamura K, Fuzzy programming with fuzzy decisions and fuzzy simulation-based genetic algorithm, *Fuzzy Sets and Systems*, Vol.122, No.2, 253-262, 2001.

[117] Liu B, Fuzzy random chance-constrained programming, *IEEE Transactions on Fuzzy Systems*, Vol.9, No.5, 713-720, 2001.

[118] Liu B, Fuzzy random dependent-chance programming, *IEEE Transactions on Fuzzy Systems*, Vol.9, No.5, 721-726, 2001.

[119] Liu B, *Theory and Practice of Uncertain Programming*, Physica-Verlag, Heidelberg, 2002.

[120] Liu B, Toward fuzzy optimization without mathematical ambiguity, *Fuzzy Optimization and Decision Making*, Vol.1, No.1, 43-63, 2002.

[121] Liu B, and Liu YK, Expected value of fuzzy variable and fuzzy expected value models, *IEEE Transactions on Fuzzy Systems*, Vol.10, No.4, 445-450, 2002.

[122] Liu B, Random fuzzy dependent-chance programming and its hybrid intelligent algorithm, *Information Sciences*, Vol.141, Nos.3-4, 259-271, 2002.

[123] Liu B, Inequalities and convergence concepts of fuzzy and rough variables, *Fuzzy Optimization and Decision Making*, Vol.2, No.2, 87-100, 2003.

[124] Liu B, *Uncertainty Theory*, Springer-Verlag, Berlin, 2004.

[125] Liu B, A survey of credibility theory, *Fuzzy Optimization and Decision Making*, Vol.5, No.4, 387-408, 2006.

[126] Liu B, A survey of entropy of fuzzy variables, *Journal of Uncertain Systems*, Vol.1, No.1, 4-13, 2007.

[127] Liu B, *Uncertainty Theory*, 2nd ed., Springer-Verlag, Berlin, 2007.

[128] Liu B, Fuzzy process, hybrid process and uncertain process, *Journal of Uncertain Systems*, Vol.2, No.1, 3-16, 2008.

[129] Liu B, *Theory and Practice of Uncertain Programming*, 2nd ed., Springer-Verlag, Berlin, 2009.

[130] Liu B, Some research problems in uncertainty theory, *Journal of Uncertain Systems*, Vol.3, No.1, 3-10, 2009.

[131] Liu B, Uncertain entailment and modus ponens in the framework of uncertain logic, *Journal of Uncertain Systems*, Vol.3, No.4, 243-251, 2009.

[132] Liu LZ, and Li YZ, The fuzzy quadratic assignment problem with penalty: New models and genetic algorithm, *Applied Mathematics and Computation*, Vol.174, No.2, 1229-1244, 2006.

[133] Liu SF, and Lin Y, *Grey Information: Theory and Practical Applications*, Springer-Verlag, London, 2006.

[134] Liu W, and Xu J, Some properties on expected value operator for uncertain variables, *Proceedings of the Eighth International Conference on Information and Management Sciences*, Kunming, China, July 20-28, 2009, pp.808-811. (Also available at http://orsc.edu.cn/online/090611.pdf)

[135] Liu XC, Entropy, distance measure and similarity measure of fuzzy sets and their relations, *Fuzzy Sets and Systems*, Vol.52, 305-318, 1992.

[136] Liu XW, Measuring the satisfaction of constraints in fuzzy linear programming, *Fuzzy Sets and Systems*, Vol.122, No.2, 263-275, 2001.

[137] Liu YH, How to generate uncertain measures, *Proceedings of Tenth National Youth Conference on Information and Management Sciences*, August 3-7, 2008, Luoyang, pp.23-26.

[138] Liu YH, and Ha MH, Expected value of function of uncertain variables, *Proceedings of the Eighth International Conference on Information and Management Sciences*, Kunming, China, July 20-28, 2009, pp.779-781. (Also available at http://orsc.edu.cn/online/090220.pdf)

[139] Liu YH, Currency models with uncertain exchange rate, http://orsc.edu.cn/online/091010.pdf.

[140] Liu YH, Uncertain stock model with periodic dividends. http://orsc.edu.cn/online/091011.pdf.

[141] Liu YK, and Liu B, Random fuzzy programming with chance measures defined by fuzzy integrals, *Mathematical and Computer Modelling*, Vol.36, Nos.4-5, 509-524, 2002.

[142] Liu YK, and Liu B, Fuzzy random variables: A scalar expected value operator, *Fuzzy Optimization and Decision Making*, Vol.2, No.2, 143-160, 2003.

[143] Liu YK, and Liu B, Expected value operator of random fuzzy variable and random fuzzy expected value models, *International Journal of Uncertainty, Fuzziness & Knowledge-Based Systems*, Vol.11, No.2, 195-215, 2003.

[144] Liu YK, and Liu B, A class of fuzzy random optimization: Expected value models, *Information Sciences*, Vol.155, Nos.1-2, 89-102, 2003.

[145] Liu YK, and Liu B, On minimum-risk problems in fuzzy random decision systems, *Computers & Operations Research*, Vol.32, No.2, 257-283, 2005.

[146] Liu YK, and Liu B, Fuzzy random programming with equilibrium chance constraints, *Information Sciences*, Vol.170, 363-395, 2005.

[147] Liu YK, Fuzzy programming with recourse, *International Journal of Uncertainty, Fuzziness & Knowledge-Based Systems*, Vol.13, No.4, 381-413, 2005.

[148] Liu YK, Convergent results about the use of fuzzy simulation in fuzzy optimization problems, *IEEE Transactions on Fuzzy Systems*, Vol.14, No.2, 295-304, 2006.

[149] Liu YK, and Wang SM, On the properties of credibility critical value functions, *Journal of Information and Computing Science*, Vol.1, No.4, 195-206, 2006.

[150] Liu YK, and Gao J, The independence of fuzzy variables with applications to fuzzy random optimization, *International Journal of Uncertainty, Fuzziness & Knowledge-Based Systems*, Vol.15, Supp.2, 1-20, 2007.

[151] Loo SG, Measures of fuzziness, *Cybernetica*, Vol.20, 201-210, 1977.

[152] Lopez-Diaz M, and Ralescu DA, Tools for fuzzy random variables: Embeddings and measurabilities, *Computational Statistics & Data Analysis*, Vol.51, No.1, 109-114, 2006.

[153] Lu M, On crisp equivalents and solutions of fuzzy programming with different chance measures, *Information: An International Journal*, Vol.6, No.2, 125-133, 2003.

[154] Lucas C, and Araabi BN, Generalization of the Dempster-Shafer Theory: A fuzzy-valued measure, *IEEE Transactions on Fuzzy Systems*, Vol.7, No.3, 255-270, 1999.

[155] Luhandjula MK, and Gupta MM, On fuzzy stochastic optimization, *Fuzzy Sets and Systems*, Vol.81, 47-55, 1996.

[156] Luhandjula MK, Fuzzy stochastic linear programming: Survey and future research directions, *European Journal of Operational Research*, Vol.174, No.3, 1353-1367, 2006.

[157] Maiti MK, and Maiti MA, Fuzzy inventory model with two warehouses under possibility constraints, *Fuzzy Sets and Systems*, Vol.157, No.1, 52-73, 2006.

[158] Maleki HR, Tata M, and Mashinchi M, Linear programming with fuzzy variables, *Fuzzy Sets and Systems*, Vol.109, No.1, 21-33, 2000.

[159] Merton RC, Theory of rational option pricing, *Bell Journal of Economics and Management Science*, Vol.4, 141-183, 1973.

[160] Mizumoto M, and Tanaka K, Some properties of fuzzy sets of type 2, *Information and Control*, Vol.31, 312-340, 1976.

[161] Molchanov IS, *Limit Theorems for Unions of Random Closed Sets*, Springer-Verlag, Berlin, 1993.

[162] Nahmias S, Fuzzy variables, *Fuzzy Sets and Systems*, Vol.1, 97-110, 1978.

[163] Negoita CV, and Ralescu D, *Simulation, Knowledge-based Computing, and Fuzzy Statistics*, Van Nostrand Reinhold, New York, 1987.

[164] Neumaier A, *Interval Methods for Systems of Equations*, Cambridge University Press, New York, 1990.

[165] Nguyen HT, On conditional possibility distributions, *Fuzzy Sets and Systems*, Vol.1, 299-309, 1978.

[166] Nguyen HT, Fuzzy sets and probability, *Fuzzy sets and Systems*, Vol.90, 129-132, 1997.

[167] Nguyen HT, Nguyen NT, and Wang TH, On capacity functionals in interval probabilities, *International Journal of Uncertainty, Fuzziness & Knowledge-Based Systems*, Vol.5, 359-377, 1997.

[168] Nguyen HT, Kreinovich V, and Wu BL, Fuzzy/probability similar to fractal/smooth, *International Journal of Uncertainty, Fuzziness & Knowledge-Based Systems*, Vol.7, 363-370, 1999.

[169] Nguyen HT, Some mathematical structures for computational information, *Information Sciences*, Vol.128, 67-89, 2000.

[170] Nguyen VH, Fuzzy stochastic goal programming problems, *European Journal of Operational Research*, Vol.176, No.1, 77-86, 2007.

[171] Nilsson NJ, Probabilistic logic, *Artificial Intelligence*, Vol.28, 71-87, 1986.

[172] Øksendal B, *Stochastic Differential Equations*, 6th ed., Springer-Verlag, Berlin, 2005.

[173] Pal NR, and Pal SK, Object background segmentation using a new definition of entropy, *IEE Proc. E*, Vol.136, 284-295, 1989.

[174] Pal NR, and Pal SK, Higher order fuzzy entropy and hybrid entropy of a set, *Information Sciences*, Vol.61, 211-231, 1992.

[175] Pal NR, and Bezdek JC, Measuring fuzzy uncertainty, *IEEE Transactions on Fuzzy Systems*, Vol.2, 107-118, 1994.

[176] Pawlak Z, Rough sets, *International Journal of Information and Computer Sciences*, Vol.11, No.5, 341-356, 1982.

[177] Pawlak Z, *Rough Sets: Theoretical Aspects of Reasoning about Data*, Kluwer, Dordrecht, 1991.

[178] Pedrycz W, Optimization schemes for decomposition of fuzzy relations, *Fuzzy Sets and Systems*, Vol.100, 301-325, 1998.

[179] Peng J, and Liu B, Stochastic goal programming models for parallel machine scheduling problems, *Asian Information-Science-Life*, Vol.1, No.3, 257-266, 2002.

[180] Peng J, and Liu B, Parallel machine scheduling models with fuzzy processing times, *Information Sciences*, Vol.166, Nos.1-4, 49-66, 2004.

[181] Peng J, and Zhao XD, Some theoretical aspects of the critical values of birandom variable, *Journal of Information and Computing Science*, Vol.1, No.4, 225-234, 2006.

[182] Peng J, and Liu B, Birandom variables and birandom programming, *Computers & Industrial Engineering*, Vol.53, No.3, 433-453, 2007.

[183] Peng J, A general stock model for fuzzy markets, *Journal of Uncertain Systems*, Vol.2, No.4, 248-254, 2008.

[184] Peng J, A stock model for uncertain markets, http://orsc.edu.cn/online/090209.pdf.

[185] Peng ZX, Some properties of product uncertain measure, http://orsc.edu.cn/online/081228.pdf.

[186] Peng ZX, and Iwamura K, A sufficient and necessary condition of uncertainty distribution, http://orsc.edu.cn/online/090305.pdf.

[187] Puri ML, and Ralescu D, Fuzzy random variables, *Journal of Mathematical Analysis and Applications*, Vol.114, 409-422, 1986.

[188] Qin ZF, and Li X, Option pricing formula for fuzzy financial market, *Journal of Uncertain Systems*, Vol.2, No.1, 17-21, 2008.

[189] Qin ZF, and Gao X, Fractional Liu process with application to finance, *Mathematical and Computer Modelling*, Vol.50, Nos.9-10, 1538-1543, 2009.

[190] Qin ZF, and Kar S, Single-period inventory problem under uncertain environment, http://orsc.edu.cn/online/090310.pdf.

[191] Qin ZF, Expected value and variance of lognormal uncertain variable, http://orsc.edu.cn/online/090420.pdf.

[192] Ralescu DA, and Sugeno M, Fuzzy integral representation, *Fuzzy Sets and Systems*, Vol.84, No.2, 127-133, 1996.

[193] Ralescu AL, and Ralescu DA, Extensions of fuzzy aggregation, *Fuzzy Sets and systems*, Vol.86, No.3, 321-330, 1997.

[194] Ramer A, Conditional possibility measures, *International Journal of Cybernetics and Systems*, Vol.20, 233-247, 1989.

[195] Ramík J, Extension principle in fuzzy optimization, *Fuzzy Sets and Systems*, Vol.19, 29-35, 1986.

[196] Saade JJ, Maximization of a function over a fuzzy domain, *Fuzzy Sets and Systems*, Vol.62, 55-70, 1994.

[197] Sakawa M, Nishizaki I, Uemura Y, Interactive fuzzy programming for two-level linear fractional programming problems with fuzzy parameters, *Fuzzy Sets and Systems*, Vol.115, 93-103, 2000.

[198] Shafer G, *A Mathematical Theory of Evidence*, Princeton University Press, Princeton, NJ, 1976.

[199] Shannon CE, *The Mathematical Theory of Communication*, The University of Illinois Press, Urbana, 1949.

[200] Shao Z, and Ji XY, Fuzzy multi-product constraint newsboy problem, *Applied Mathematics and Computation*, Vol.180, No.1, 7-15, 2006.

[201] Shih HS, Lai YJ, and Lee ES, Fuzzy approach for multilevel programming problems, *Computers and Operations Research,* Vol.23, 73-91, 1996.

[202] Shreve SE, *Stochastic Calculus for Finance II: Continuous-Time Models*, Springer, Berlin, 2004.

[203] Slowinski R, and Teghem J, Fuzzy versus stochastic approaches to multicriteria linear programming under uncertainty, *Naval Research Logistics*, Vol.35, 673-695, 1988.

[204] Slowinski R, and Vanderpooten D, A generalized definition of rough approximations based on similarity, *IEEE Transactions on Knowledge and Data Engineering*, Vol.12, No.2, 331-336, 2000.

[205] Steuer RE, Algorithm for linear programming problems with interval objective function coefficients, *Mathematics of Operational Research*, Vol.6, 333-348, 1981.

[206] Sugeno M, *Theory of Fuzzy Integrals and its Applications*, Ph.D. Dissertation, Tokyo Institute of Technology, 1974.

[207] Szmidt E, Kacprzyk J, Entropy for intuitionistic fuzzy sets, *Fuzzy Sets and Systems*, Vol.118, 467-477, 2001.

[208] Taleizadeh AA, Niaki STA, and Aryanezhad MB, A hybrid method of Pareto, TOPSIS and genetic algorithm to optimize multi-product multi-constraint inventory control systems with random fuzzy replenishments, *Mathematical and Computer Modelling*, Vol.49, Nos.5-6, 1044-1057, 2009.

[209] Tanaka H, and Asai K, Fuzzy solutions in fuzzy linear programming problems, *IEEE Transactions on Systems, Man and Cybernetics*, Vol.14, 325-328, 1984.

[210] Tanaka H, and Guo P, *Possibilistic Data Analysis for Operations Research*, Physica-Verlag, Heidelberg, 1999.

[211] Torabi H, Davvaz B, Behboodian J, Fuzzy random events in incomplete probability models, *Journal of Intelligent & Fuzzy Systems*, Vol.17, No.2, 183-188, 2006.

[212] Wang G, and Liu B, New theorems for fuzzy sequence convergence, *Proceedings of the Second International Conference on Information and Management Sciences*, Chengdu, China, August 24-30, 2003, pp.100-105.

[213] Wang GJ, and Wang H, Non-fuzzy versions of fuzzy reasoning in classical logics, *Information Sciences*, Vol.138, No.1-4, 211-236, 2001.

[214] Wang GJ, Formalized theory of general fuzzy reasoning, *Information Sciences*, Vol.160, No.1-4, 251-266, 2004.

[215] Wang XZ, Wang YD, and Xu XF, A new approach to fuzzy rule generation: fuzzy extension matrix, *Fuzzy Sets and Systems*, Vol.123, No.3, 291-306, 2001.

[216] Wang ZY, and Klir GJ, *Fuzzy Measure Theory*, Plenum Press, New York, 1992.

[217] Yager RR, On measures of fuzziness and negation, Part I: Membership in the unit interval, *International Journal of General Systems*, Vol.5, 221-229, 1979.

[218] Yager RR, Entropy and specificity in a mathematical theory of evidence, *International Journal of General Systems*, Vol.9, 249-260, 1983.

[219] Yager RR, On the entropy of fuzzy measures, *IEEE Transactions on Fuzzy Systems*, Vol.8, 453-461, 2000.

[220] Yager RR, On the evaluation of uncertain courses of action, *Fuzzy Optimization and Decision Making*, Vol.1, 13-41, 2002.

[221] Yang L, and Liu B, On inequalities and critical values of fuzzy random variable, *International Journal of Uncertainty, Fuzziness & Knowledge-Based Systems*, Vol.13, No.2, 163-175, 2005.

[222] Yang L, and Liu B, On continuity theorem for characteristic function of fuzzy variable, *Journal of Intelligent & Fuzzy Systems*, Vol.17, No.3, 325-332, 2006.

[223] Yang L, and Li K, B-valued fuzzy variable, *Soft Computing*, Vol.12, No.11, 1081-1088, 2008.

[224] Yang N, and Wen FS, A chance constrained programming approach to transmission system expansion planning, *Electric Power Systems Research*, Vol.75, Nos.2-3, 171-177, 2005.

[225] Yang XH, Moments and tails inequality within the framework of uncertainty theory, *Proceedings of the Eighth International Conference on Information and Management Sciences*, Kunming, China, July 20-28, 2009, pp.812-814. (Also available at http://orsc.edu.cn/online/090612.pdf)

[226] Yazenin AV, On the problem of possibilistic optimization, *Fuzzy Sets and Systems*, Vol.81, 133-140, 1996.

[227] You C, and Wen M, The entropy of fuzzy vectors, *Computers & Mathematics with Applications*, Vol.56, No.6, 1626-1633, 2008.

[228] You C, Some convergence theorems of uncertain sequences, *Mathematical and Computer Modelling*, Vol.49, Nos.3-4, 482-487, 2009.

[229] Zadeh LA, Fuzzy sets, *Information and Control*, Vol.8, 338-353, 1965.

[230] Zadeh LA, Outline of a new approach to the analysis of complex systems and decision processes, *IEEE Transactions on Systems, Man and Cybernetics*, Vol.3, 28-44, 1973.

[231] Zadeh LA, The concept of a linguistic variable and its application to approximate reasoning, *Information Sciences*, Vol.8, 199-251, 1975.

[232] Zadeh LA, Fuzzy sets as a basis for a theory of possibility, *Fuzzy Sets and Systems*, Vol.1, 3-28, 1978.

[233] Zadeh LA, A theory of approximate reasoning, In: J Hayes, D Michie and RM Thrall, eds, *Mathematical Frontiers of the Social and Policy Sciences*, Westview Press, Boulder, Cororado, 69-129, 1979.

[234] Zhao R and Liu B, Stochastic programming models for general redundancy optimization problems, *IEEE Transactions on Reliability*, Vol.52, No.2, 181-191, 2003.

[235] Zhao R, and Liu B, Renewal process with fuzzy interarrival times and rewards, *International Journal of Uncertainty, Fuzziness & Knowledge-Based Systems*, Vol.11, No.5, 573-586, 2003.

[236] Zhao R, and Liu B, Redundancy optimization problems with uncertainty of combining randomness and fuzziness, *European Journal of Operational Research*, Vol.157, No.3, 716-735, 2004.

[237] Zhao R, and Liu B, Standby redundancy optimization problems with fuzzy lifetimes, *Computers & Industrial Engineering*, Vol.49, No.2, 318-338, 2005.

[238] Zhao R, Tang WS, and Yun HL, Random fuzzy renewal process, *European Journal of Operational Research*, Vol.169, No.1, 189-201, 2006.

[239] Zhao R, and Tang WS, Some properties of fuzzy random renewal process, *IEEE Transactions on Fuzzy Systems*, Vol.14, No.2, 173-179, 2006.

[240] Zheng Y, and Liu B, Fuzzy vehicle routing model with credibility measure and its hybrid intelligent algorithm, *Applied Mathematics and Computation*, Vol.176, No.2, 673-683, 2006.

[241] Zhou J, and Liu B, New stochastic models for capacitated location-allocation problem, *Computers & Industrial Engineering*, Vol.45, No.1, 111-125, 2003.

[242] Zhou J, and Liu B, Analysis and algorithms of bifuzzy systems, *International Journal of Uncertainty, Fuzziness & Knowledge-Based Systems*, Vol.12, No.3, 357-376, 2004.

[243] Zhou J, and Liu B, Modeling capacitated location-allocation problem with fuzzy demands, *Computers & Industrial Engineering*, Vol.53, No.3, 454-468, 2007.

[244] Zhu Y, and Liu B, Continuity theorems and chance distribution of random fuzzy variable, *Proceedings of the Royal Society of London Series A*, Vol.460, 2505-2519, 2004.

[245] Zhu Y, and Liu B, Some inequalities of random fuzzy variables with application to moment convergence, *Computers & Mathematics with Applications*, Vol.50, Nos.5-6, 719-727, 2005.

[246] Zhu Y, and Ji XY, Expected values of functions of fuzzy variables, *Journal of Intelligent & Fuzzy Systems*, Vol.17, No.5, 471-478, 2006.

[247] Zhu Y, and Liu B, Fourier spectrum of credibility distribution for fuzzy variables, *International Journal of General Systems*, Vol.36, No.1, 111-123, 2007.

[248] Zhu Y, and Liu B, A sufficient and necessary condition for chance distribution of random fuzzy variables, *International Journal of Uncertainty, Fuzziness & Knowledge-Based Systems*, Vol.15, Supp.2, 21-28, 2007.

[249] Zhu Y, Fuzzy optimal control with application to portfolio selection, http://orsc.edu.cn/process/080117.pdf.

[250] Zhu Y, Uncertain optimal control with application to a portfolio selection model, http://orsc.edu.cn/online/090524.pdf.

[251] Zimmermann HJ, *Fuzzy Set Theory and its Applications*, Kluwer Academic Publishers, Boston, 1985.

[252] Zuo Y, and Ji XY, Theoretical foundation of uncertain dominance, *Proceedings of the Eighth International Conference on Information and Management Sciences*, Kunming, China, July 20-28, 2009, pp.827-832. (Also available at http://orsc.edu.cn/online/081231.pdf)

[253] Zuo Y, Critical values of function of uncertain variables, http://orsc.edu.cn/online/090807.pdf.

List of Frequently Used Symbols

$$\mathcal{M} \quad \text{uncertain measure}$$

$\mathcal{M}$	uncertain measure
$(\Gamma, \mathcal{L}, \mathcal{M})$	uncertainty space
$\xi,\, \eta,\, \tau$	uncertain variables
$\Phi,\, \Psi,\, \Upsilon$	uncertainty distributions
$\lambda,\, \rho,\, \mu,\, \nu$	identification functions
$\mathcal{L}(a,b)$	linear uncertain variable
$\mathcal{Z}(a,b,c)$	zigzag uncertain variable
$\mathcal{N}(e,\sigma)$	normal uncertain variable
$\mathcal{LOGN}(e,\sigma)$	lognormal uncertain variable
E	expected value
V	variance
H	entropy
$\xi_{\sup}(\alpha)$	α-optimistic value to ξ
$\xi_{\inf}(\alpha)$	α-pessimistic value to ξ
X_t, Y_t, Z_t	uncertain processes
C_t	canonical process
N_t	renewal process
$\emptyset$	empty set
$\Re$	set of real numbers
$\Re^n$	set of n-dimensional real vectors
$\vee$	maximum operator
$\wedge$	minimum operator
$\rhd$	matching operator
$a_i \uparrow a$	$a_1 \le a_2 \le \cdots$ and $a_i \to a$
$a_i \downarrow a$	$a_1 \ge a_2 \ge \cdots$ and $a_i \to a$
$A_i \uparrow A$	$A_1 \subset A_2 \subset \cdots$ and $A = A_1 \cup A_2 \cup \cdots$
$A_i \downarrow A$	$A_1 \supset A_2 \supset \cdots$ and $A = A_1 \cap A_2 \cap \cdots$

Index

Baoding Liu
Uncertainty Theory

Uncertainty theory is a branch of mathematics based on normality, monotonicity, self-duality, countable subadditivity, and product measure axioms. Uncertainty is any concept that satisfies the axioms of uncertainty theory. The goal of uncertainty theory is to study the behavior of uncertain phenomena. This book provides a self-contained, comprehensive and up-to-date presentation of uncertainty theory, including uncertain programming, uncertain process, uncertain calculus, uncertain differential equation, uncertain logic, uncertain entailment, and uncertain inference. Mathematicians, researchers, engineers, designers, and students in the field of mathematics, information science, operations research, system science, industrial engineering, computer science, artificial intelligence, finance, control, and management science will find this work a stimulating and useful reference.

Axiom 1. (Normality Axiom) $\mathcal{M}\{\Gamma\} = 1$ for the universal set Γ.

Axiom 2. (Monotonicity Axiom) $\mathcal{M}\{\Lambda_1\} \leq \mathcal{M}\{\Lambda_2\}$ whenever $\Lambda_1 \subset \Lambda_2$.

Axiom 3. (Self-Duality Axiom) $\mathcal{M}\{\Lambda\} + \mathcal{M}\{\Lambda^c\} = 1$ for any event Λ.

Axiom 4. (Countable Subadditivity Axiom) For every countable sequence of events $\Lambda_1, \Lambda_2, \cdots$, we have

$$\mathcal{M}\left\{\bigcup_{i=1}^{\infty} \Lambda_i\right\} \leq \sum_{i=1}^{\infty} \mathcal{M}\{\Lambda_i\}.$$

Axiom 5. (Product Measure Axiom) Let $(\Gamma_k, \mathcal{L}_k, \mathcal{M}_k)$ be uncertainty spaces for $k = 1, 2, \cdots, n$. The product uncertain measure $\mathcal{M} = \mathcal{M}_1 \wedge \mathcal{M}_2 \wedge \cdots \wedge \mathcal{M}_n$ is

$$\mathcal{M}\{\Lambda\} = \begin{cases} \sup\limits_{\Lambda_1 \times \Lambda_2 \times \cdots \times \Lambda_n \subset \Lambda} \min\limits_{1 \leq k \leq n} \mathcal{M}_k\{\Lambda_k\}, \\ \qquad \text{if} \sup\limits_{\Lambda_1 \times \Lambda_2 \times \cdots \times \Lambda_n \subset \Lambda} \min\limits_{1 \leq k \leq n} \mathcal{M}_k\{\Lambda_k\} > 0.5 \\ 1 - \sup\limits_{\Lambda_1 \times \Lambda_2 \times \cdots \times \Lambda_n \subset \Lambda^c} \min\limits_{1 \leq k \leq n} \mathcal{M}_k\{\Lambda_k\}, \\ \qquad \text{if} \sup\limits_{\Lambda_1 \times \Lambda_2 \times \cdots \times \Lambda_n \subset \Lambda^c} \min\limits_{1 \leq k \leq n} \mathcal{M}_k\{\Lambda_k\} > 0.5 \\ 0.5, \quad \text{otherwise.} \end{cases}$$

Principle. (Maximum Uncertainty Principle) For any event, if there are multiple reasonable values that an uncertain measure may take, then the value as close to 0.5 as possible is assigned to the event.