Cui Yanhong

Guo Renkuan

Dunne Tim
University of Cape Town, Cape Town, South Africa

Guo Danni
South African National Biodiversity Institute, Cape Town, South Africa

Decision theory under general uncertainty

Keywords

statistical decision, state space, action space, uncertain measure, loss function, risk, uncertainty decision

Abstract

The exposure of Toyota management's cover-up of its faulty car component problems raises a fundamental question: did Toyota management make an appropriate decision taking all uncertainties into account? Statistical decision theory is a framework with a probabilistic foundation, which admits random uncertainty about the real world and human thinking. In general, the uncertainty of the real world is diversified and therefore the effort of trying to deal with different forms of uncertainty with one special form of uncertainty, namely random uncertainty, may be oversimplified. In this paper, we introduce an axiomatic uncertain measure theoretical framework and explore the essential mechanism in formulating a general uncertainty decision theory. We expect that a new understanding of uncertainty and development of a corresponding new uncertainty decision-making approach may assist intelligence communities to survive and deal with the extremely tough and diverse aspects of an uncertain reality.

1. Introduction

Recently a "Made in Japan" crisis began spreading widely, triggered by a Toyota Prius brake fault, a fire-sparking Honda Jazz electric-window, and a Sony Camera problem, and a cover-up of faulty seats for Boeing 747 jets. These events have shocked worldwide business and industry.

Journalism today has largely castigated the Toyota decision makers for censoring and even hiding fault factors from the public, particularly, attempting to prevent release to the press. In some sense, the journalists are correct, *a posteriori*, but in the decision point of view, they are not necessarily correct *a priori*, because not all the "factors" (states of nature) of car making necessarily enter into the decision mechanisms – the model distributions, as when some states are allocated a tiny possibility or ignored. Furthermore, active members of safety and reliability communities may do well to calmly re-examine whether or not the theoretical roots and foundation of statistical decision theory are sufficient for the purpose to which it is applied. Randomness is merely one of the forms of uncertainty. In consequence, existing statistical theory may not always provide a completely suitable analysis for data embodying more general uncertainty.

The real world is not as simple as we imagine. Uncertainty is intrinsic and diversified in form. For example, vagueness is a different form of uncertainty from randomness, and enters more and more into today's industrial environments, as Carvalho and Machado [1] have commented, "In a global market, companies must deal with a high rate of changes in business environment. … The parameters, variables and restrictions of the production system are inherently vagueness." Therefore one may argue that a company decision should no longer only be a somewhat routine exercise of applying traditional statistical decision techniques arising from its sound but constrained probabilistic underpinnings. Without a thorough and explicit exploration of uncertainty and its characteristics, attempts to abstract real world uncertainty into appropriate concepts will inevitably permit that decision exercises to fall short of the reality of business. Our “reality”, its diversity, and the formality of a general uncertainty are the fundamental rationale for us to pursue the exploration of a new general uncertainty decision approach.

Decision making is based on a consensus “truth”, whether or not that truth exists in the eyes of other communities. Furthermore, some factors, such as the faulty seat phenomena and cover-up are not repeatable incidents, and cannot admit complete probability assignment, i.e., assigning probability to such an isolated incident is illogical, but, top managements must include such “accident” factors in decision-making. Thus, it is logical to say decision is a subjective activity.

The discipline of Statistics builds upon probability theory and deals with collecting, analysing, and drawing conclusions from data information essentially featured by random uncertainty and imposed upon modelled pattern (in terms of a probability distribution). It is essential to emphasize that the mechanism underlying statistical decision is the use of probability distributions. An uncertainty decision problem is essentially the appropriate specification of uncertain distributions.

The fundamental problem here is what uncertainty distributions one may invoke to characterize the relevant states and events. Recently, Professor Baoding Liu from Tsinghua University (Beijing, China) proposed an axiomatic uncertainty measure theory [6], which is sub- σ- additive and less restrictive than the (σ- additive) probability measure. A σ -algebra, denoted by $\mathfrak{A}(\Xi)$ is a collection of subsets (events) in a set Ξ satisfying three properties: (i) $\Xi \in \mathfrak{A}(\Xi)$; (ii) $\forall A \in \mathfrak{A}(\Xi)$, $A^c \in \mathfrak{A}(\Xi)$, where A^c is the complement of A in Ξ ; (iii) $\bigcup_{n=1}^{+\infty} A_n \in \mathfrak{A}(\Xi)$ Probability measure and uncertain measure may be defined on a σ -algebra, and each characterises a probability distribution or a uncertainty distribution respectively. By virtue of its less restrictive nature, Liu’s [6] uncertain measure theory can support uncertainty distribution building, analysis and modelling more general uncertainty observations and their use in the making of a decision.

The remainder of the paper is structured as follows: Section two introduces Liu’s [1] new axiomatic uncertain measure theory, which permits the new uncertainty distribution underlying the decision mechanism; Section three reviews the (probabilistic) statistical decision theory in order to reveal the underlying mechanism behind statistical decision making – relevant (probability) distributions. This review suggests that the new general uncertainty decision approach should preserve the basic framework of the (probabilistic) statistical decision theory but in some conditions replace the underlying probability distributions with appropriate uncertain distributions. Section four discusses the basic elements and some intrinsic features of uncertainty decision theory in comparisons with statistical decision theory. Sections five and six illustrate the general uncertainty decision making approach in discrete and continuous uncertainty environments respectively; Section seven concludes the paper.

2. Uncertain measure foundation

Uncertain measure [6] is an axiomatically defined set function mapping from a σ -algebra of a given space (set) to the unit interval [0,1], which provides a measuring grade system for an uncertain phenomenon and permits the formal definition of an uncertain variable.

Let Ξ be a nonempty set (space), and $\mathfrak{A}(\Xi)$ the σ -algebra on Ξ. Each subset $A \subset \Xi$, $A \in \mathfrak{A}(\Xi)$ is called an uncertain event. A number denoted $\lambda\{A\}$, $0 \le \lambda\{A\} \le 1$, is assigned to event A, which indicates the uncertain measuring grade with which event A occurs. Occurrence of an event A is defined as occurrence of any constituent outcome *x* within A . The set function $\lambda\{A\}$ satisfies the following axioms given by Liu [6]:

Axiom 1: (Normality) $\lambda\{\Xi\} = 1$.

Axiom 2: (Monotonicity) $\lambda\{\cdot\}$ is non-decreasing: If $A \subset B$, then $\lambda\{A\} \le \lambda\{B\}$.

Axiom 3: (Self-Duality) $\lambda\{\cdot\}$ is self-dual, If $A \in \mathfrak{A}(\Xi)$, then $\lambda\{A\} + \lambda\{A^c\} = 1$, where A^c is the complement of A in Ξ.

Axiom 4: (σ- Subadditivity) $\lambda\left\{\bigcup_{i=1}^{\infty} A_i\right\} \le \sum_{i=1}^{\infty} \lambda\{A_i\}$ for any countable event sequence $\{A_i\}$.

*Definition 2.1***.** (Liu [6]) Any set function $\lambda : \mathfrak{A}(\Xi) \to [0,1]$ satisfying *Axioms L.1-L.4* is called an uncertain measure. The triple $(\Xi, \mathfrak{A}(\Xi), \lambda)$ is called an uncertain space.

Definition 2.2. [6] An uncertain variable ξ is a measurable mapping, i.e., $\xi:(\Xi,\mathfrak{A}(\Xi))\to(\mathbb{R},\mathfrak{B}(\mathbb{R}))$, where $\mathfrak{B}(\mathbb{R})$ is the Borel σ- algebra on $\mathbb{R}=(-\infty,+\infty)$.

Remark 2.3. The Borel σ- algebra is the smallest set class of Borel sets on $\mathbb{R}=(-\infty,+\infty)$. The Borel sets include $\mathbb{R}=(-\infty,+\infty)$, empty set $\varnothing$, all the closed intervals, $[a,b]$, all the semi-closed intervals, $[a,b)$, $(a,b]$, $[a,+\infty)$, and $(-\infty,b]$, and all the open intervals, (a,b), $(-\infty,b)$, and $(a,+\infty)$, where $\forall a,b\in\mathbb{R}$, $-\infty<a\le b<+\infty$.

Remark 2.4. The fundamental difference between a random variable and an uncertain variable is the measure space on which they are defined. In the associated triples, the first two elements are similar in form: the set and a σ- algebra on the set. However, the third elements in the triples: the measures defined on the σ- -algebras, are not similar. The former (i.e. the probability measure Pr) obeys σ- additivity and the later (i.e. the uncertain measure λ obeys only σ-subadditivity. The choice of a measure inevitably has impacts on the behaviour of any measurable function on the triple.

Definition 2.5. [6] The uncertain distribution $\Psi:\mathbb{R}\to[0,1]$ of an uncertain variable $\xi:\Xi\to\mathbb{R}$ defined on the uncertain space $\Xi,\mathfrak{A}\ \Xi\ ,\lambda$ is

$$\Psi(x)=\lambda\{\tau\in\Xi|\xi(\tau)\le x\} \tag{1}$$

Remark 2.6. A random variable X is a measurable mapping. To understand the measurability of a random variable, particularly, the role played by the σ-algebra $\mathfrak{F}(\Omega)$, we note how measurability is structured for a random variable. Let $(\Omega,\mathfrak{F},P)$ be a probability space and $(\mathbb{R},\mathfrak{B}(\mathbb{R}))$ be a measurable space on real-line, then a real-valued function X is random variable if and only if the pre-image $\{\omega\in\Omega:X(\omega)\le r\}\in\mathfrak{F}$, for all $r\in\mathbb{R}$,. For each value $r\in\mathbb{R}$ taken by a real-valued random variable X, the event $B=(-\infty,r]$ is an element of the Borel σ-algebra over a real-line $\mathbb{R}$, the pre-image of event B under random variable X is an event

$$\omega\in\Omega: X\ \ \omega\ \in B\ =\ \omega\in\Omega: X\ \ \omega\ \le r \tag{2}$$

where $\{\omega\in\Omega:X(\omega)\le r\}$ is an element of σ- algebra $\mathfrak{F}$ over Ω, and the probability measure P is defined on this set class, i.e., σ- algebra $\mathfrak{F}$, i.e., $P:\mathfrak{F}\to[0,1]$. Therefore every element (event) of $\mathfrak{F}$ is assigned with a probability grade, i.e., event $\{\omega\in\Omega:X(\omega)\le r\}$ is assigned a probability grade, which is $P\{\omega\in\Omega:X(\omega)\le r\}$.

Thus, σ-algebra $\mathfrak{F}$ facilitates the formal definition of a random variable in terms of the membership of the pre-image $\{\omega\in\Omega:X(\omega)\le r\}$ within the σ- algebra $\mathfrak{F}$, on which the probability measuring grade defined. Every event in σ -algebra $\mathfrak{F}$ is assigned a probability. Each random variable on the probability space $(\Omega,\mathfrak{F},P)$ induces a probability space $(\mathbb{R},\mathfrak{B}(\mathbb{R}),\mu)$ by means of the following well-known correspondence.

$$\forall B\in\mathfrak{B}\ \mathbb{R}\ :\ \mu\ B\ =P\ \ X^{-1}\ B\ \ =P\ \ X\in B \tag{3}$$

Let us write $\mu=P\circ X^{-1}$ and specifically, the probability distribution is defined by the induced measure μ,

$$F\ \ r\ =\mu\ \ -\infty,r\ \ =P\ \ X\le r \tag{4}$$

In all, the random variable X defined on a given probability space $(\Omega,\mathfrak{F}(\Omega),P)$ is a measurable mapping to $(\mathbb{R},\mathfrak{B}(\mathbb{R}))$ and thus induces the distribution function, $F:\mathbb{R}\to[0,1]$, which is used to characterize the random variable.

Similarly, in the axiomatic development of uncertain measure, the σ -algebra $\mathfrak{A}(\Xi)$ plays critical roles as the set class in defining both the measurability of an uncertain variable ξ and the set function λ as the uncertain measure. The roles are equivalent to the roles played by a σ-algebra in probability measure theory in defining both the measurability of a random variable X and the set function P as the probability measure. As long as an uncertain measure λ is specified, the uncertain distribution Ψ is fully defined. The next theorem states the necessary and sufficiency conditions for a function to be an uncertain distribution.

Theorem 2.7 [7] Let $\Psi:\mathbb{R}\to[0,1]$ be a non-decreasing function with

$$\Psi(-\infty)=0,\Psi(+\infty)=1. \tag{5}$$

Then set function $\nu:\mathfrak{B}(\mathbb{R})\to[0,1]$, for any Borel set B:

$$\nu\{B\}=\begin{cases}\nu^*\{B\} & \text{if } \nu^*\{B\}<0.5\\ 1-\nu^*\{B^c\} & \text{if } \nu^*\{B^c\}<0.5\\ 0.5 & \text{otherwise}\end{cases} \tag{6}$$

where

$$\nu^*\{B\}=\nu_1\{B\}\wedge\nu_2\{B\}\wedge\nu_3\{B\},\forall B\in\mathfrak{B}(\mathbb{R}) \tag{7}$$

with $\nu_i:\mathfrak{B}(\mathbb{R})\to[0,1]$, $i=1,2,3$, given by

$$\nu_1\{B\}=\begin{cases}1-\lim\limits_{x\uparrow\inf\{B\}}\Psi(x) & \text{if } \inf\{B\}\in B\\ 1-\Psi(\inf\{B\}) & \text{otherwise}\end{cases}, \tag{8}$$

$$\nu_2\{B\}=\Psi(\sup\{B\}), \tag{9}$$

and

$$\nu_3\{B\}=\inf_{(a,b]\subset B^c}\{\Psi(a)+1-\Psi(b)\}, \tag{10}$$

is an uncertain measure on the Borel σ-algebra, $\mathfrak{B}(\mathbb{R})$.

Remark 2.8. Technically, once the induced uncertain measure ν is defined by Eq. (6), the uncertain space $(\mathbb{R},\mathfrak{B}(\mathbb{R}),\nu)$ is established. Then the mapping from $(\mathbb{R},\mathfrak{B}(\mathbb{R}),\nu)$ to $(\Xi,\mathfrak{A}(\Xi),\lambda)$ can be built, i.e., the uncertain measure λ is fully specified.
For comparison purposes, we note the definition of the probability distribution.

Definition 2.9. Let $(\Omega,\mathfrak{F}(\Omega),P)$ be a given probability space, the probability distribution of a random variable on $(\Omega,\mathfrak{F}(\Omega),P)$ is

$$F(x)=P\{\omega\in\Omega|X(\omega)\le x\} \tag{11}$$

Theorem 2.10. Let $F:\mathbb{R}\to[0,1]$. Then, F is a probability distribution function if and only if F satisfies each of the following three conditions:
(i) $\lim_{x\to-\infty}F(x)=0,\ \lim_{x\to+\infty}F(x)=1$;
(ii) $F(x)$ is non-decreasing in x;
(iii) F is right-continuous, i.e., $\forall x_0\in\mathbb{R}$, $\lim_{x\downarrow x_0}F(x)=F(x_0)$.

Remark 2.11. The difference between a probability distribution and an uncertainty distribution relates to whether the distribution possesses right-continuity. The relaxation of the condition in the uncertain distribution function arises from the sub-σ-additivity property of the underlying uncertain measure λ. On the basis of this distributional difference, the new uncertainty decision theory developed in this paper differs from statistical decision theory without relying on any arguments about any interpretation differences of the two distributions.

Definition 2.12. (Liu [6]) An *n*-dimensional uncertain vector from an uncertain measure space $(\Xi,\mathfrak{A}(\Xi),\lambda)$ to the set of *n*-dimensional real-valued vectors, i.e., for Borel set B within $\mathbb{R}^n$, the set

$$\{\underline{\xi}\in B\}=\{\tau\in\Xi|\underline{\xi}(\tau)\in B\} \tag{12}$$

is an event.

Theorem 2.13. [6] Let $\underline{\xi}=(\xi_1,\xi_2,\cdots,\xi_n)^T$ be an uncertain vector, and $f:\mathbb{R}^n\to\mathbb{R}$ a measurable function. Then $f(\underline{\xi})$ is an uncertain variable such that

$$\lambda\{f(\underline{\xi})\in B\}=\lambda\{\underline{\xi}\in f^{-1}(B)\} \tag{13}$$

for any Borel set B within $\mathbb{R}^n$.

Now, we are ready to investigate statistical decision theory and extend its principles to uncertainty decision theory because we have adequate self-contained materials to understand the further explorations.

3. Statistical decision theory

Statistical decision theory is established on the axiomatic foundation of probability measure, see [4] and [8]. The developments can be sourced in [2] and [3]. A measure theoretical decision theory is stated below using well-known results.

3.1 Three elements of statistical decision

The three elements of statistical decision theory are: the states of the nature, the action space, and the loss in elementary statistical course. At the measure theoretical level, in order to simplify the mechanism underlying statistical decision, we use "sample space" instead of "states of nature". Note here the term "statistical decision" implies that a decision is made by observation-based statistical analysis. Without data and the distribution underlying the data, there is nothing. While both "sample space and distributional family" and "states of nature" may share the same meanings, the former is more comprehensive and characteristic-exposing and the latter is more intuitive. In discussing statistical decision theory we use "sample space and distributional family" to emphasize the observational and data oriented statistical nature of the set for specifying a decision problem.

(1) *Sample space and distributional family*. In measure theoretical language, every concrete value of X, denoted as x, is called a sample value. The set of all possible sample values contribute to a sample space, denoted as $\mathfrak{X} = \{x, x \in X\}$. It is necessary to emphasize that the specification of $\mathfrak{X}$ only requires $\mathfrak{X}$ contains all the possible values of X, but it is not required that X must admit all values in $\mathfrak{X}$. For example, though X may take non-negative real values, i.e., $X = \mathbb{R}^+ = [0, +\infty)$ we may nonetheless introduce the sample space as $\mathfrak{X} = \mathbb{R} = (-\infty, +\infty)$. This property may provide great conveniences in mathematical treatments later. Furthermore, a σ-algebra (field) is specified on $\mathfrak{X}$, denoted as $\mathfrak{B}(\mathfrak{X})$. In statistical inferences, in specifying sample space, it is necessary to specify both $\mathfrak{X}$ and $\mathfrak{B}(\mathfrak{X})$, i.e., the measurable space $(\mathfrak{X}, \mathfrak{B}(\mathfrak{X}))$. Thus, in measure theoretical treatments, it is often the practice to regard the measurable space $(\mathfrak{X}, \mathfrak{B}(\mathfrak{X}))$ as a sample space, instead of simply $\mathfrak{X}$ as in elemental statistics. The most commonly used sample space $\mathfrak{X}$ is the *n*-dimensional Euclidean space $\mathbb{R}^n$, while $\mathfrak{B}(\mathfrak{X})$ is the σ-algebra of the Borel sets in $\mathbb{R}^n$, or $\mathfrak{X}$ is a subset of the *n*-dimensional Euclidean space $\mathbb{R}^n$, while $\mathfrak{B}(\mathfrak{X})$ is the σ-algebra of the Borel sets of $\mathfrak{X}$. Often there is no rigorous distinction between these two forms. Without special claims, when using $(\mathfrak{X}, \mathfrak{B}(\mathfrak{X})) = (\mathbb{R}^n, \mathfrak{B}(\mathbb{R}^n))$, the later cases are covered. $(\mathbb{R}^n, \mathfrak{B}(\mathbb{R}^n))$ is called the Euclidean sample space.

On the σ-algebra $\mathfrak{B}(\mathfrak{X})$, a family of probability measures $\{P_\theta, \theta \in \Theta\}$ is defined, where Θ is called a parameter space, and in many situations, $\Theta \subset \mathbb{R}^m$. The distribution of X is one of the members in the distributional family $\{P_\theta, \theta \in \Theta\}$, i.e., there exists a $\theta_0 \in \Theta$, such that the distribution of X is P_{θ_0} but the value of θ_0 is unknown. Determining a value of θ_0, i.e., the specific distribution for X from the distributional family $\{P_\theta, \theta \in \Theta\}$ is precisely the object of the statistical inference.

The sample space and distributional family together determine the probability mechanism of the observations (i.e., sample values) from the population X. This pair is often written in the form of $\{(\mathfrak{X}, \mathfrak{B}(\mathfrak{X}), P_\theta), \theta \in \Theta\}$ or alternatively, we say that the sample space of X is $(\mathfrak{X}, \mathfrak{B}(\mathfrak{X}))$ with distributional family $\{P_\theta, \theta \in \Theta\}$.
Recall that $(\Omega, \mathfrak{A}, P)$ is called a probability space, therefore, it is important to bring the family of the probability spaces $\{(\mathfrak{X}, \mathfrak{B}(\mathfrak{X}), P_\theta), \theta \in \Theta\}$ into the decision process since the selection of θ is the basic implicit task.

(2) *Decision space*. Statistical decision making (or inference), whether it takes the form of point estimation, or interval estimation, or hypothesis testing, is actually decision making based on the sample information (statistics). The set of all possible decision outcomes constitutes of a decision space, denoted by $\mathfrak{D}$. For the requirements of the measure theoretical developments, a σ-algebra on $\mathfrak{D}$ is necessary, denoted by $\mathfrak{B}(\mathfrak{D})$. Thus $(\mathfrak{D}, \mathfrak{B}(\mathfrak{D}))$ is called as a decision space.

(3) *Loss function*. Whether a decision is bad or good, there must be a platform for comparison. For given probability distributional family $\{(\mathfrak{X}, \mathfrak{B}_\mathfrak{X}, P_\theta), \theta \in \Theta\}$ and the decision space $\mathfrak{D}$, we introduce a loss function associated with a specific decision d, $d \in \mathfrak{D}$, for the justification of decision merit.

Definition 3.1. Let the parameter space be Θ and the decision space be $(\mathfrak{D}, \mathfrak{B}(\mathfrak{D}))$. Any function defined on the product space $\Theta \times \mathfrak{D}$ is a loss function, denoted as $L(\theta, d)$, if it is satisfies the following two conditions:
(a) $0 \le L(\theta, d) < \infty$ for any $\theta \in \Theta$ and any $d \in \mathfrak{D}$;
(b) for any fixed $\theta \in \Theta$, $L(\theta, d)$ as function of d is $\mathfrak{B}(\mathfrak{D})$-measurable.
The specification of loss sets up the criterion for decision choices, namely control of loss.

3.2 Decision function

Having the descriptions of the three basic elements of statistical decision problem, we note that for any concrete problem, decision making aims to select a good decision d in $\mathfrak{D}$, which depends upon the value of loss function $L(\theta,d)$. If θ is given, the problem is easily settled. Given a value of θ, the distribution of X, P_θ, is known. If θ is not given, it is necessary to utilize the information of the observational data and the underlying distribution of X contained in the sample values x (to infer the value θ) in order to support the decision maker to make the choice.

Therefore, the task of a statistical decision is just to establish a function $\delta(x)$ called a statistical decision function, defined on the sample space $(\mathfrak{X},\mathfrak{B}(\mathfrak{X}))$ and taking values on decision space $(\mathfrak{D},\mathfrak{B}(\mathfrak{D}))$ such that when a sample value x is available, the value of the decision $\delta(x)$ will be determined.

Definition 3.2. (Non-randomized statistical decision function) Let the sample space be $(\mathfrak{X},\mathfrak{B}(\mathfrak{X}))$ and the decision space be $(\mathfrak{D},\mathfrak{B}(\mathfrak{D}))$. Any measurable transformation $\delta(x)$ defined on $\mathfrak{X}$ and taking values on $\mathfrak{D}$ is termed a non-randomized statistical decision function.

For any individual decision problem, there are many possible decision functions available. It is necessary to introduce a numerical index, the risk function, to reflect the quality of decision function $\delta(x)$.

Definition 3.3. (Risk function) Suppose that the sample space and distributional family is given by $(\mathfrak{X},\mathfrak{B}(\mathfrak{X}),P_\theta)$, $\theta\in\Theta$, the decision space is $(\mathfrak{D},\mathfrak{B}(\mathfrak{D}))$, the loss function is $L(\theta,d)$, and $\delta(x)$ is a decision function. A function of θ, called the risk function denoted as $R(\theta,\delta)$ is defined with respect to a decision function δ, as

$$\begin{aligned} R(\theta,\delta) &= \mathrm{E}_\theta\left[L(\theta,\delta(X))\right] \\ &= \int_{\mathfrak{X}} L(\theta,\delta(x))\,dP_\theta(x), \quad \theta\in\Theta \end{aligned} \tag{14}$$

In other words, a risk function is the average loss if the decision $\delta(x)$ is taken for whatever the random value x is observed, when the true parameter θ is given (or assumed). It is obvious that the lesser the risk, the better the decision.

Definition 3.4. (Randomized statistical decision function) Let the sample space be $(\mathfrak{X},\mathfrak{B}(\mathfrak{X}))$ and the decision space be $(\mathfrak{D},\mathfrak{B}(\mathfrak{D}))$. A function defined on the space $\mathfrak{X}\times\mathfrak{B}(\mathfrak{D})$, $\delta(x,D)$, is a randomized statistical decision function, if

(a) for any fixed $D\in\mathfrak{B}_{\mathfrak{D}}$, $\delta(x,D)$ as the function of x is $\mathfrak{B}(\mathfrak{X})$-measurable;

(b) for any fixed $x\in\mathfrak{X}$, $\delta(x,D)$ as the function of D is a probability measure on $\mathfrak{B}(\mathfrak{D})$.

In adopting a randomized statistical decision function δ, the procedure for obtaining a decision d is as follows: first, obtain a random sample x by observing the population X, then in terms of δ, obtain the probability measure $\delta(x,D)$ on $\mathfrak{B}(\mathfrak{D})$, and finally, in terms of the probability measure $\delta(x,D)$, select a decision d from the decision space $\mathfrak{D}$. It is obvious that the previously defined non randomized statistical decision function is a special case of the randomized statistical decision function defined here, i.e., for any $x\in\mathfrak{X}$, probability distribution $\delta(x,D)$ is concentrated at a point in D as the function of D (Clearly this point is partially determined by x, *i.e.,* $\delta(x)$ *in Definition 3.2.*)

One might query the significance of introducing such a very abstract and seemingly unnatural concept. At this stage, we simply stress that the concept will bring certain theoretical conveniences.

The risk function of a randomized statistical decision function is described as follows. Let the sample space and the distributional family be $\{(\mathfrak{X},\mathfrak{B}(\mathfrak{X}),P_\theta),\theta\in\Theta\}$, decision space be $(\mathfrak{D},\mathfrak{B}(\mathfrak{D}))$, and the loss function be $L(\theta,d)$. In order to calculate the risk function $R(\theta,\delta)$ for the randomized statistical decision function we set $\delta=\delta(x,D)$. It is easy to establish that given the sample x, the conditional risk of δ is

$$R(\theta,\delta\,|\,x)=\int_{\mathfrak{X}} L(\theta,\omega)\delta(x,d\omega) \tag{15}$$

where ω is the moving point in $\mathfrak{D}$. It is necessary to note that

$$\begin{aligned} R(\theta,\delta\,|\,x) &= \int_{\mathfrak{X}} L(\theta,\omega)\delta(x,d\omega) \\ &= \int_{\mathfrak{X}} dP_\theta\left[\int_{\mathfrak{D}} L(\theta,\omega)\delta(x,d\omega)\right] \end{aligned} \tag{16}$$

In order to make the definition effective, it is necessary to verify: the integrals in (2) and (3) are meaningful. Note that the functions involved are all nonnegative. Hence once the measurability of the functions is confirmed, the conclusion is reached. $L(\theta,\omega)$ as a function of ω is $\mathfrak{B}(\mathfrak{D})$-measurable. Now assuming the integral in (2) to be meaningful, then the integral in (3) must be meaningful. This confirmation arises from showing that $R(\theta,\delta\,|\,x)$ as a function of x, is $\mathfrak{B}(\mathfrak{X})$-measurable. In other words, for any nonnegative $\mathfrak{B}(\mathfrak{D})$-measurable function $f(\omega)$, the function

$$\gamma(x) = \int_{\mathfrak{D}} f(\omega)\delta(x, d\omega) \tag{17}$$

must be $\mathfrak{B}(\mathfrak{X})$-measurable. This property can be shown by the standard measure theoretical approach: start with indicator function $f(\omega) = \vartheta_D(\omega)$, $D \in \mathfrak{B}(\mathfrak{D})$, then a simple function $f(\omega) = \sum_{i=1}^{n} \vartheta_{D_i}(\omega),\ \ D_i \in \mathfrak{B}(\mathfrak{D})$, and finally an arbitrary nonnegative function $f(\omega)$, which can be approached by limiting process by specifying a sequence of nonnegative simple functions $\{f_n(\omega)\}$.

4. Elements of uncertainty decision problem

Now, with the descriptions of the statistical decision theory, we may attempt to discuss the construction of a general uncertainty decision theory in a comparable manner.

It is necessary to point out the three components, i.e., the *state, action, and loss* in the statistical decision theory, are still the essential elements in the new general uncertainty decision theory. (Note, the term for the first element, “state” is used here). However, the connotations inherent in the three elements are not always the same as in the statistical approach. Let us examine element by element.

Firstly, in statistical decision theory, the state, termed “state of nature” (i.e., sample space and distributional family), is regarded as objectively in existence, at least in some consensus sense, while in any general uncertainty environments, the state may include subjective judgmental or even phenomenological events or factors. For example, top decision makers may include company’s middle managements’ or engineers possible cover-up behaviour as one of the “state” elements, which need not be observable and non-repeatable events. (Such possible information may sound dirty, spurious or problematic, and the decision makers might never wish to release this approach to the employees or the public). Note here the conceptual interpretations of state acquire when involving the decision environments, i.e., “reality” ahead of the decision makers, possible virtual actions, and virtual loss. The differentiation between the “state of nature” in the statistical decision theory and the “state” in the uncertainty decision theory is critical. The former is under the frequentist statistician’s “reality” inferences, more or less reflecting the “truth”, while the latter is a mixture of subjective and objective reflections.

Secondly, the connotations of “action” in the statistical decision theory and that in the uncertainty decision theory can be discerned. The action in the statistical decision theory is the possible reaction and treatment against the state event. The state element is the root cause and accordingly, the action is selected against the cause. It is an if-then logic and thus a many to one mapping from the state space onto action space (point-wise).

While the “action” in the uncertainty decision theory is not necessarily a one-to-one mapping consequence because particular “states” under consideration may be of artificial, or personal experience, or a phenomenological nature. A director with rich experience may not be afraid in facing a crisis and s/he may even delicately utilize the crisis to create more development chances instead. The state space in an experienced director’s mind is likely to be much enlarged by inclusion of many “states” due to specific understanding of crisis, in contrast with an inexperienced junior manager, Therefore, the action space is virtual, in which some elements are of a precautionary nature and do not correspond to any specific state element. The mapping is of multiple states to multiple action nature. However, the inclusion of virtual action elements is extremely important, because the top decision maker does not need to deal with routine decisions of day-to-day operations but with the extreme event or the most important event decision.

Thirdly, the loss in both decision theories is the same. However, the social loss and environmental loss occupy more and more concern from the public, NGO and the governmental agencies as well. In the new uncertainty decision theory, safety factor state, health factor state, and environmental factor state should be automatically assigned uncertain measure grades because of their intrinsic features.

According to statistical decision theory, the decision is made in terms of observational data, denoted as z, which is described by an probability distribution $F(z;\theta)$. Based on data z (i.e., $F(z;\theta)$), a decision is actually a mapping from data space $\mathbb{D}$ into action space $\mathbb{A}$. In other words,

$$a:\mathbb{D}\to\mathbb{A} \tag{18}$$

which can be expressed by

$$a=d(z) \tag{19}$$

The action taken is random because the observational data to which decision rule is applied are random. Consequently, the loss function $l(\theta,d(z))$ is random in statistical decision theory. Therefore, $\{(\mathfrak{X},\mathfrak{B}(\mathfrak{X}),P_\theta),\theta\in\Theta\}$, the probability distribution family is constituted of the elementary mechanism underlying the statistical decision processes.

Similarly, in section two, where the uncertainty distribution theory is introduced, we stress that $\{(\mathbb{Z},\mathfrak{B}(\mathbb{Z}),\Psi_\tau),\tau\in\Gamma\}$, the uncertainty distribution family is constituted of the fundamental mechanism underlying the uncertainty decision processes. Let us explore what an uncertainty distribution may look like via a detailed example.

Example 4.1. Let ξ be an uncertain variable, which takes values on $\mathbb{Z}=\{0,0+,1,1+,2,2+,3,3+,4\}$ with the uncertainty distribution:

$$\Psi_\xi(z)=\begin{cases}0 & z<0\\ 0.25 & z=0\\ 0.2z+0.25 & 0<z<1\\ 0.575 & z=1\\ 0.125z+0.45 & 1<z<2\\ 0.77 & z=2\\ 0.07z+0.63 & 2<z<3\\ 0.85 & z=3\\ 0.09z+0.59 & 3<z<4\\ 1.0 & z\ge 4\end{cases} \tag{20}$$

which implies

$$\begin{aligned}
&\pi_0=\lambda\{\xi=0\}=0.25,\\
&\pi_{0+}=\lambda\{\xi=0+\}=0.05,\\
&\pi_{1-}=\lambda\{\xi=1-0\}=0.15,\\
&\pi_1=\lambda\{\xi=1\}=0.125,\\
&\pi_{1+}=\lambda\{\xi=1+0\}=0.015,\\
&\pi_{2-}=\lambda\{\xi=2-0\}=0.11,\\
&\pi_2=\lambda\{\xi=2\}=0.07,\\
&\pi_{2+}=\lambda\{\xi=2+0\}=0.03,\\
&\pi_{3-}=\lambda\{\xi=3-0\}=0.04,\\
&\pi_3=\lambda\{\xi=3\}=0.01,\\
&\pi_{3+}=\lambda\{\xi=3+0\}=0.08,\\
&\pi_{4-}=\lambda\{\xi=4-0\}=0.02,\\
&\pi_4=\lambda\{\xi=4-0\}=0.05
\end{aligned} \tag{21}$$

Remark 4.2. It is easily seen that the uncertain distribution Ψ defined in Example 4.1 is a function with finite jumps, where $\lim_{z\uparrow z_i}\xi=c_{i-}<c_{i+}=\lim_{z\downarrow z_i}\xi$ and $\xi=z_i=c_i$, $c_{i-}<c_i<c_{i+}$, satisfying *Theorem 2.7*, although $\{z=c_i\}$, $i=0,1,\cdots,4$ are so-called removable points in calculus theory. It can be further verified that the uncertain distribution Ψ is neither left continuous nor right-continuous. Ψ in this example gives an elementary form of an uncertain distribution.

Definition 4.3. (Essential Form of an Uncertain Distribution) Let ξ be an uncertain variable with essential form, which takes its values from an ascending ordered domain set $\mathbb{D}=\{c_0,c_1,\cdots,c_n\}$ with the uncertain distribution Ψ defined by

$$\begin{aligned}
&\Psi(c_0)=\lambda\{\xi\le c_0\}=0,\\
&\Psi(c_1-)=\psi_{1-},\ \Psi(c_1)=\psi_1,\ \Psi(c_1+)=\psi_{1+},\\
&\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots,\\
&\Psi(c_i-)=\psi_{i-},\ \Psi(c_i)=\psi_i,\ \Psi(c_i+)=\psi_{i+},\\
&\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots,\\
&\Psi(c_n-)=\psi_{n-},\Psi(c_n)=1.00
\end{aligned} \tag{22}$$

such that $\psi_{i-}<\psi_i<\psi_{i+}$, $i=1,2,\cdots,n$. Furthermore, it requires $\pi_i=\lambda\{\xi=c_i\}\in(0,\psi_{i+}-\psi_{i-})$, $i=1,2,\cdots,n$.

Theorem 4.4. Let ξ be an uncertain variable with the essential form, which takes values from ascending ordered domain set $\mathbb{D}=\{c_0,c_1,\cdots,c_n\}$ with the uncertain distribution. Then Ψ satisfies the following necessary and sufficient conditions:

(i)

$$\Psi(c_0)=\lambda\{\xi\le c_0\}=0 \tag{23}$$

(ii) For $i=1,2,\cdots,n-1$,

$$\begin{aligned} &\Psi(c_i -) = \text{ƛ}\{\xi < c_i\} = \psi_{i-}, \\ &\Psi(\{c_i\}) = \text{ƛ}\{\xi = c_i\} = \pi_i, \\ &\bar{\Psi}(c_i +) = \text{ƛ}\{\xi > c_i\} = \psi_{i+} \end{aligned} \tag{24}$$

(iii)

$$\begin{aligned} &\Psi(c_n -) = \text{ƛ}\{\xi < c_n\} = \psi_{n-}, \\ &\Psi(\{c_n\}) = \text{ƛ}\{\xi = c_n\} = \pi_n, \\ &\Psi(c_n) = \text{ƛ}\{\xi \le c_n\} = 1.00 \end{aligned} \tag{25}$$

(iv) The uncertain measure of singleton $\{c_i\}$

$$\begin{aligned} &\pi_{i-} = \text{ƛ}\{\xi = c_i - 0\}, \\ &\pi_i = \text{ƛ}\{\xi = c_i\}, \\ &\pi_{i+} = \text{ƛ}\{\xi = c_i + 0\}, \end{aligned} \tag{26}$$

such that $\sum_{i=0}^{n}(\pi_{i-} + \pi_i + \pi_{i+}) = 1$.

Definition 4.5. If an uncertainty distribution takes the form

$$\Psi_d(z) = \begin{cases} 0 & z < c_0 \\ \pi_1 & z = c_1 \\ \pi_{1+} & c_1 < z < c_2 \\ \pi_2 & z = c_2 \\ \vdots & \vdots \\ \pi_i & z = c_i \\ \pi_{i+} & c_i < z < c_{i+1} \\ \pi_{i+1} & z = c_{i+1} \\ \vdots & \vdots \\ 1.0 & z \ge c_m \end{cases} \tag{27}$$

where $0 < \pi_i < \pi_{i+} < \pi_{i+1} < 1$, then it is a discrete uncertain distribution.

Theorem 4.6. The expectation of a discrete uncertainty distribution Ψ_d, denoted as $\mathrm{E}_\Psi[\xi]$, is given by

$$\mathrm{E}_\Psi[\xi] = \sum_{i=0}^{n} w_i c_i \tag{28}$$

where

$$\begin{aligned} w_i = &\max_{0 \le j \le n}\{\pi_j, \pi_{j+} \mid c_j \le c_i\} \wedge 0.5 \\ &- \max_{0 \le j \le n}\{\pi_j, \pi_{j+} \mid c_j < c_i\} \wedge 0.5 \\ &+ \max_{0 \le j \le n}\{\pi_j, \pi_{j+} \mid c_j \ge c_i\} \wedge 0.5 \\ &- \max_{0 \le j \le n}\{\pi_j, \pi_{j+} \mid c_j > c_i\} \wedge 0.5 \end{aligned} \tag{29}$$

$i = 0, 1, 2, \cdots, m$.

Proof: The proof of Theorem 6.4 is just the application of Liu's [6] definition of uncertain expectation to a discrete uncertain variable with neither left-continuity nor right-continuity:

$$\mathrm{E}[\xi] = \int_0^{+\infty} \text{ƛ}\{\xi \ge s\}\, ds - \int_{-\infty}^{0} \text{ƛ}\{\xi \le s\}\, ds. \tag{30}$$

Example 4.7. Calculate the expectation of the discrete uncertain variable defined by

$$\Psi_\xi(z) = \begin{cases} 0 & z < 0 \\ 0.25 & z = 0 \\ 0.45 & 0 < z < 1 \\ 0.575 & z = 1 \\ 0.7 & 1 < z < 2 \\ 0.77 & z = 2 \\ 0.84 & 2 < z < 3 \\ 0.85 & z = 3 \\ 0.95 & 3 < z < 4 \\ 1.0 & z \ge 4 \end{cases} \tag{31}$$

Let us calculate the weight w_i, $i = 0, 1, 2, 3, 4$. Note that the uncertain measure grades can be written

$$\begin{aligned} &\pi_0 = \text{ƛ}\{\xi = 0\} = 0.25, \\ &\pi_{0+} = \text{ƛ}\{\xi = 0+\} = 0.20, \\ &\pi_1 = \text{ƛ}\{\xi = 1\} = 0.125, \\ &\pi_{1+} = \text{ƛ}\{\xi = 1+0\} = 0.125, \\ &\pi_2 = \text{ƛ}\{\xi = 2\} = 0.07, \\ &\pi_{2+} = \text{ƛ}\{\xi = 2+0\} = 0.07, \\ &\pi_3 = \text{ƛ}\{\xi = 3\} = 0.01, \\ &\pi_{3+} = \text{ƛ}\{\xi = 3+\} = 0.10, \\ &\pi_4 = \text{ƛ}\{\xi = 4-0\} = 0.05 \end{aligned} \tag{32}$$

Then we can calculate the weights:

$$\begin{aligned} w_0 = &\max_{0 \le j \le n}\{\pi_j \mid j \le 0\} \wedge 0.5 - \max_{1 \le j \le n}\{\pi_j \mid j < 0\} \wedge 0.5 \\ &+ \max_{1 \le j \le n}\{\pi_{j+}, \pi_j \mid j \ge 0\} \wedge 0.5 - \max_{1 \le j \le n}\{\pi_{j+}, \pi_j \mid j > 0\} \wedge 0.5 \\ = &\,0.25 - 0.00 + 0.20 - 0.125 = 0.325 \end{aligned}$$

$$\begin{aligned} w_1 = &\max_{0 \le j \le n}\{\pi_j, \pi_{j+} \mid j \le 1\} \wedge 0.5 - \max_{0 \le j \le n}\{\pi_j, \pi_{j+} \mid j < 1\} \wedge 0.5 \\ &+ \max_{0 \le j \le n}\{\pi_j, \pi_{j+} \mid j \ge 1\} \wedge 0.5 - \max_{0 \le j \le n}\{\pi_j, \pi_{j+} \mid j > 1\} \wedge 0.5 \\ = &\,0.25 - 0.20 + 0.125 - 0.125 = 0.05 \end{aligned}$$

$$\begin{aligned} w_2 = &\max_{0 \le j \le n}\{\pi_{j-}, \pi_j, \pi_{j+} \mid j \le 2\} \wedge 0.5 - \max_{0 \le j \le n}\{\pi_{j-}, \pi_j, \pi_{j+} \mid j < 2\} \wedge 0.5 \\ &+ \max_{0 \le j \le n}\{\pi_{j-}, \pi_j, \pi_{j+} \mid j \ge 2\} \wedge 0.5 - \max_{0 \le j \le n}\{\pi_{j-}, \pi_j, \pi_{j+} \mid j > 2\} \wedge 0.5 \\ = &\,0.25 - 0.25 + 0.10 - 0.10 = 0.00 \end{aligned}$$

$$\begin{aligned} w_3 = &\max_{0 \le j \le n}\{\pi_{j-}, \pi_j, \pi_{j+} \mid j \le 3\} \wedge 0.5 - \max_{0 \le j \le n}\{\pi_{j-}, \pi_j, \pi_{j+} \mid j < 3\} \wedge 0.5 \\ &+ \max_{0 \le j \le n}\{\pi_{j-}, \pi_j, \pi_{j+} \mid j \ge 3\} \wedge 0.5 - \max_{0 \le j \le n}\{\pi_{j-}, \pi_j, \pi_{j+} \mid j > 3\} \wedge 0.5 \\ = &\,0.25 - 0.25 + 0.10 - 0.10 = 0.00 \end{aligned}$$

and

$$\begin{aligned} w_4 = &\max_{0 \le j \le n}\{\pi_{j-}, \pi_j, \pi_{j+} \mid j \le 4\} \wedge 0.5 - \max_{0 \le j \le n}\{\pi_{j-}, \pi_j, \pi_{j+} \mid j < 4\} \wedge 0.5 \\ &+ \max_{0 \le j \le n}\{\pi_{j-}, \pi_j, \pi_{j+} \mid j \ge 4\} \wedge 0.5 - \max_{0 \le j \le n}\{\pi_{j-}, \pi_j, \pi_{j+} \mid j > 4\} \wedge 0.5 \\ = &\,0.25 - 0.25 + 0.05 - 0.00 = 0.05 \end{aligned}$$

Hence the expected value of the uncertain variable is

$$\mathrm{E}[\xi] = 0.25\times 0 + 0.05\times 1 + 0.00\times 2 + 0.00\times 3 + 0.05\times 4 = 0.25$$

Remark 4.8. Recall that in statistical decision, the index, risk, is an expectation involving a distribution function. In other words, the decision element - loss function is a random quantity in statistical decision theory, or it is an uncertain quantity in uncertainty decision theory. It is also well-known fact that we cannot compare two random quantities directly until the statistical experiment is finished and the outcomes are obtained we do not know their values. Furthermore, because each individual outcome (realization) of a random variable (quantity) is associated with a probability, simply comparing the values of two random quantities will let decision-making miss their associated probability grades. Therefore, statistical decision is commonly made on the weighted average sense, i.e., expectation sense. The risk concept is just reflecting such a feature. Similarly, it is fair to say the uncertainty decision is also made using expectation of an uncertainty distribution.

The fact that the Toyota management was aware of faulty car component problems, but did not address potential impacts of social risk and environmental awareness, may reveal that the their description of the state space is questionable. In other words, the state space Θ used in the Toyota decision might include the faulty parts sub-space $\Theta_f \subset \Theta$, but the associated measure grades assigned to the states in Θ_f were so tiny so that the decision function specification essentially generates near-null loss, and hence decisions unaffected by the parts are preferable to the management's thinking.

According to Liu [6], the uncertain distribution function on the state space is a prior distribution, $p(\theta)$, which may or may not be updateable by the information in data form. Bayes theorem (in probability) may also provide a updating structure for the state distribution, whose result is the called posterior uncertain distribution on the state $p(\theta \mid z)$.

A fundamental issue here is the uncertain prior distribution specification, particularly, the uncertain measure grades in Θ_f .

5. A discrete uncertainty decision

Definition 5.1. An uncertain decision is a selection, which minimizes the loss function $l(\theta,a)$ or regret function $r(\theta,a) = l(\theta,a) - \min_{a\in\mathbb{A}}\{l(\theta,a)\}$ of an action a from action space $\mathbb{A}$ for given state θ in the state space Θ .

Definition 5.2. The expected value of the loss with respect to the distribution of uncertain data z .

$$R(\theta,d) = \mathrm{E}_{\theta}\left[l(\theta,d(z))\right] \quad (33)$$

is called the risk function.

Remark 5.3. The distribution of uncertain data z depends on state θ , because the dependence of $R(\theta,d)$ on θ enters explicitly from $l(\theta,a)$ and also through the state θ in the distribution function $\Psi(z;\theta)$ for z .

Example 5.4. Two quality states $\Theta = \{\theta_1,\theta_2\}$, where

θ_1: Liu's [6] uncertain quality state, θ_2, Gaussian quality state are assumed.

The uncertain variable ξ is discrete variable taking values on $\mathbb{Z} = \{0, 0+, 1, 1+, 2, 2+, 3, 3+, 4\}$. Then $\Psi(z;\theta)$, the uncertain distribution given θ_1 is:

$$\Psi_{\xi}(z \mid \theta_1) = \begin{cases} 0 & z<0 \\ 0.71 & z=0 \\ 0.71964 & 0<z<1 \\ 0.93964 & z=1 \\ 0.94822 & 1<z<2 \\ 0.99122 & z=2 \\ 0.99138 & 2<z<3 \\ 0.99938 & z=3 \\ 0.99998 & 3<z<4 \\ 1.0 & z\ge 4 \end{cases} \quad (34)$$

Note that the uncertain measure grades given θ_1 are

$$\begin{aligned} \pi_0 &= \lambda\{\xi=0\} = 0.71, \\ \pi_{0+} &= \lambda\{\xi=0+\} = 0.00964, \\ \pi_1 &= \lambda\{\xi=1\} = 0.22, \\ \pi_{1+} &= \lambda\{\xi=1+0\} = 0.00858, \\ \pi_2 &= \lambda\{\xi=2\} = 0.04, \\ \pi_{2+} &= \lambda\{\xi=2+0\} = 0.0032, \\ \pi_3 &= \lambda\{\xi=3\} = 0.008, \\ \pi_{3+} &= \lambda\{\xi=3+\} = 0.0006, \\ \pi_4 &= \lambda\{\xi=4\} = 0.00002 \end{aligned} \quad (35)$$

The uncertain distribution $\Psi(z;\theta)$ given θ_2 is:

$$\Psi_\xi(z \mid \theta_2)=\begin{cases} 0 & z<0 \\ 0.68 & z=0 \\ 0.68268 & 0<z<1 \\ 0.95268 & z=1 \\ 0.9545 & 1<z<2 \\ 0.9965 & z=2 \\ 0.9973 & 2<z<3 \\ 0.9993 & z=3 \\ 0.99995 & 3<z<4 \\ 1.0 & z\geq 4 \end{cases} \tag{36}$$

Note that the uncertain measure grades are

$$\begin{aligned} &\pi_0 = ƛ\{\xi=0\}=0.68, \\ &\pi_{0+} = ƛ\{\xi=0+\}=0.00268, \\ &\pi_1 = ƛ\{\xi=1\}=0.27, \\ &\pi_{1+} = ƛ\{\xi=1+0\}=0.00182, \\ &\pi_2 = ƛ\{\xi=2\}=0.042, \\ &\pi_{2+} = ƛ\{\xi=2+0\}=0.0008, \\ &\pi_3 = ƛ\{\xi=3\}=0.002, \\ &\pi_{3+} = ƛ\{\xi=3+\}=0.00065, \\ &\pi_4 = ƛ\{\xi=4\}=0.00005 \end{aligned} \tag{37}$$

The loss is defined in the *Table 1*. The uncertain distribution is defined by *Table 2.*

Table 1. A loss function in tabular form

$l(\theta,a)$	θ_1	θ_2
a_1	1.2	4.5
a_2	3.5	1.5

Table 2. Data uncertain distribution

$\Pr\{Z=z_i \mid \theta_j\}$	θ_1	θ_2
$\xi=z_1=0$	0.71	0.68
$\xi=z_2=0+$	0.00964	0.00268
$\xi=z_3=1$	0.22	0.27
$\xi=z_4=1+$	0.00858	0.00182
$\xi=z_5=2$	0.004	0.042
$\xi=z_6=2+$	0.0032	0.0008
$\xi=z_7=3$	0.008	0.002
$\xi=z_8=3+$	0.0006	0.00065
$\xi=z_9=4$	0.00002	0.00005

Then the decision function will have $9\times 2=18$ elements since there are 2 actions and 9 observations :

Table 3. Uncertain decision function $a=d(z)$

$d(z)$	d_1	d_2	d_3	d_4	d_5	d_6	d_7	d_8	d_9
z_1	a_1	a_1	a_1	a_1	a_1	a_1	a_1	a_1	a_2
z_2	a_1	a_1	a_1	a_1	a_1	a_1	a_1	a_2	a_2
z_3	a_1	a_1	a_1	a_1	a_1	a_1	a_2	a_2	a_2
z_4	a_1	a_1	a_1	a_1	a_1	a_2	a_2	a_2	a_2
z_5	a_1	a_1	a_1	a_1	a_2	a_2	a_2	a_2	a_2
z_6	a_1	a_1	a_1	a_2	a_2	a_2	a_2	a_2	a_2
z_7	a_1	a_1	a_2	a_2	a_2	a_2	a_2	a_2	a_2
z_8	a_1	a_2	a_2	a_2	a_2	a_2	a_2	a_2	a_2
z_9	a_2	a_2	a_2	a_2	a_2	a_2	a_2	a_2	a_2

where the decision space $\mathbb{D}=\{d_1,d_2,\cdots,d_9\}$.

Table 4. Risk function $R\ \theta,d\ = \mathrm{E}_\theta\left[l\ \theta,d\ z\ \ \right]$

$R(\theta,d)$	θ_1	θ_2
d_1	1.200066	3.4999
d_2	1.202046	3.4986
d_3	1.228446	3.4946
d_4	1.228974	3.493
d_5	1.370874	3.40885
d_6	1.68108	3.40536
d_7	2.125188	2.88036
d_8	2.157	2.86
d_9	4.5	1.5

To demonstrate the calculation of entries in *Table 4*, we evaluate $R\ \theta_i,d_j\ ,i=1,2;\ j=1,2,\cdots,9$,

$$\begin{aligned} &R\ \theta_1,d_1\ = \\ &=l\ \theta_1,a_1\ \ p\ \ z=z_1 \mid \theta_1\ +l\ \theta_1,a_1\ \ p\ \ z=z_2 \mid \theta_1 \\ &+l\ \theta_1,a_1\ \ p\ \ z=z_3 \mid \theta_1\ +l\ \theta_1,a_1\ \ p\ \ z=z_4 \mid \theta_1 \\ &+l\ \theta_1,a_1\ \ p\ \ z=z_5 \mid \theta_1\ +l\ \theta_1,a_1\ \ p\ \ z=z_6 \mid \theta_1 \\ &+l\ \theta_1,a_1\ \ p\ \ z=z_7 \mid \theta_1\ +l\ \theta_1,a_1\ \ p\ \ z=z_8 \mid \theta_1 \\ &+l\ \theta_1,a_2\ \ p\ \ z=z_9 \mid \theta_1 \\ &=1\times 0.99998+4\times 0.00002=1.00006 \end{aligned}$$

$$\begin{aligned}&R(\theta_1,d_2)\\&=l(\theta_1,a_1)p(z=z_1|\theta_1)+l(\theta_1,a_1)p(z=z_2|\theta_1)\\&+l(\theta_1,a_1)p(z=z_3|\theta_1)+l(\theta_1,a_1)p(z=z_4|\theta_1)\\&+l(\theta_1,a_1)p(z=z_5|\theta_1)+l(\theta_1,a_1)p(z=z_6|\theta_1)\\&+l(\theta_1,a_1)p(z=z_7|\theta_1)+l(\theta_1,a_2)p(z=z_8|\theta_1)\\&+l(\theta_1,a_2)p(z=z_9|\theta_1)\\&=1\times0.99938+4\times0.00062=1.00186\end{aligned}$$

With the specifications of risk function, a further criterion for selecting a decision must be defined. As an example, we employ the *minimax principle*. The first step is to maximise the risk with respect to state space, i.e.,

$$d^M=\max_\theta\{R(\theta,d)\} \tag{32}$$

And the second step is to find the minimax decision rule which minimizes d^M

$$d^{mM}=\min_d\{d^M\}=\min_d\left\{\max_\theta\{R(\theta,d)\}\right\} \tag{33}$$

Example 5.5. Minimax decision rule for *Example 5.4.*

Table 5. Minimax decision rule d^{mM} search

	θ_1	θ_2	$M(d)$
d_1	1.200066	3.4999	3.4999
d_2	1.202046	3.4986	3.4986
d_3	1.228446	3.4946	3.4946
d_4	1.228974	3.493	3.493
d_5	1.370874	3.40885	3.40885
d_6	1.68108	3.40536	3.40536
d_7	2.125188	2.88036	2.88036
d_8	2.157	2.86	2.86
d_9	4.5	1.5	4.5
d^{mM}			**2.86**

The development of Section 5 reveals that in a discrete uncertainty decision, the procedure for the uncertainty decision is similar to that for statistical decision. The fundamental difference lies in the connotation of the "state" and underlying uncertainty distribution as discussed in Section 4.

6. A continuous uncertain decision

In Section 5 we explored the uncertain decision given discrete state space, action space and discrete loss function environments. Now, we investigate the decision problem under a continuous uncertainty environment.

Recall that *Definition 4.3* states the essential form of the uncertainty distribution. It is neither left-continuous nor right-continuous and its distribution function has finite jumps and "removable" values at jump points. If the distribution function is continuous everywhere, i.e., there is no jump and no removable point in its domain, it is a continuous uncertainty distribution.

Peng and Iwamura [7] give an uncertain variable $\xi(x)=x$ defined on the uncertain space $(\Xi,\mathfrak{A},\lambda)$ where

$$\lambda(A)=\begin{cases}0 & \text{if } A=\varnothing\\ c & \text{if } A \text{ is upper bounded}\\ 0.5 & \text{if } A \text{ and } A^c \text{ are both upper bounded}\\ 1-c & \text{if } A^c \text{ is upper bounded}\\ 1 & \text{if } A=\mathbb{R}\end{cases} \tag{34}$$

Then the uncertain distribution for ξ is $\Psi(x)=c,\ 0<c<0.5$.

Another continuous uncertainty distribution example is Liu's [6] uncertain normal distribution

$$\Psi(z;e,\sigma)=\frac{1}{1+e^{-\frac{\pi}{\sqrt{3}\sigma}(z-e)}},\quad z\in\mathbb{R} \tag{35}$$

Let us consider the uncertain decision problem for a given continuous distribution.

Assume state space $\Theta=\mathbb{R}$, and action space $\mathbb{A}=\mathbb{R}$, and the loss function defined by

$$l(\theta,a)=w(\theta)(\theta-a)^2 \tag{35}$$

i.e., a quadratic loss function is assumed.

Definition 6.1. (Uncertain Bayes loss) Given a continuous state space Θ, the uncertain variable θ is defined on uncertain space $(\Theta,\mathfrak{B}(\Theta),\lambda)$, where $\lambda(\cdot)$ is an uncertain measure. The uncertain distribution $\Psi(\theta)$ is defined on $(\Theta,\mathfrak{B}(\Theta))$). Then we seek the average of loss with respect to state space for a given action $a\in\mathbb{A}$. (Note the action space $\mathbb{A}$ is continuous too.) The quantity

$$B(a)=\mathrm{E}[l(\theta,a)]=\int_\Theta l(\theta,a)d\Psi(\theta) \tag{36}$$

is called as the uncertain Bayes loss for a given action a.

Definition 6.2. (Uncertain Bayes risk) The Bayes risk is

$$B(d) = \mathrm{E}\left[l(\theta, d(z))\right] = \int_{\Theta} l(\theta, d(z))\, d\Psi(\theta) \tag{37}$$

Definition 6.3. (Uncertain Bayes rule) A Bayes decision rule, denoted as d^B is a rule such that the Bayes risk is minimized, i.e.,

$$B(d^B) = \min_{d \in \mathbb{D}} \{B(d)\} \tag{38}$$

Example 6.4. Given a continuous state space $\Theta = \mathbb{R}$, the uncertain variable θ is defined on uncertain space $(\mathbb{R}, \mathfrak{B}(\mathbb{R}), \lambda)$, where $\lambda(\cdot)$ is properly defined. The uncertain distribution is

$$\Psi(\theta) = \begin{cases} 0 & \theta \le \alpha \\ \dfrac{\theta - \alpha}{2(\beta - \alpha)} & \alpha < \theta \le \beta \\ \dfrac{\theta + \gamma - 2\beta}{2(\gamma - \beta)} & \beta < \theta \le \gamma \\ 1 & \theta > \gamma \end{cases} \tag{39}$$

Then we find the average of loss with respect to state space for a given action $a \in \mathbb{A}$, as the uncertain Bayes loss:

$$B(a) = \mathrm{E}\left[l(\theta, a)\right] = \int_{\Theta} w(\theta)(\theta - a)^2\, d\Psi(\theta) \tag{40}$$

Set $w(\theta) = w_0$, a constant. The uncertain Bayes loss is

$$\begin{aligned} B(a) &= \frac{w_0}{2(\beta-\alpha)} \int_{\alpha}^{\beta} (\theta - a)^2 d\theta + \frac{w_0}{2(\gamma-\beta)} \int_{\beta}^{\gamma} (\theta - a)^2 d\theta \\ &= \frac{w_0}{6}\left[(\beta-a)^2 + (\alpha-a)^2 + (\beta-a)(\alpha-a)\right] \\ &+ \frac{w_0}{6}\left[(\gamma-a)^2 + (\beta-a)^2 + (\gamma-a)(\beta-a)\right] \\ &= \frac{w_0}{6}\left(3a^2 - 3(\alpha+\beta)a + \alpha^2 + \beta^2\right) \\ &+ \frac{w_0}{6}\left(3a^2 - 3(\gamma+\beta)a + \gamma^2 + \beta^2\right) \\ &= \frac{w_0}{6}\left(6a^2 - 3(\alpha + 2\beta + \gamma)a + \left(\alpha^2 + 2\beta^2 + \gamma^2\right)\right) \end{aligned} \tag{41}$$

With appropriate specification of decision function in term *of* data, the uncertain Bayesian decision analysis can be formulated.

Liu [6] states his maximum uncertain principle, (abbreviated as MUP): *"for any event, if there are multiple reasonable values that an uncertain measure may take, then the value as close as to 0.5 as possible is assigned to the event".*

Definition 6.5. Let $\omega(\cdot \mid x)$ denote the regular conditional distribution of θ, given $X = x$. If $\{v_\theta, \theta \in \Theta\} \ll \mu$, where μ is a σ-finite measure on $\mathfrak{B}_{\mathfrak{X}}$, and define $f(\theta, x) = dv_\theta(x)/d\mu$, then for $\forall B \in \mathfrak{B}_\Theta$,

$$\omega(B \mid x) = \begin{cases} \dfrac{\int_B f(\theta,x)\, d\omega(\theta)}{\int_\Theta f(\theta,x)\, d\omega(\theta)} \vee 0.5 & \text{if } \dfrac{\int_B f(\theta,x)\, d\omega(\theta)}{\int_\Theta f(\theta,x)\, d\omega(\theta)} < 0.5 \\ 1 - \dfrac{\int_{B^c} f(\theta,x)\, d\omega(\theta)}{\int_\Theta f(\theta,x)\, d\omega(\theta)} & \text{if } \dfrac{\int_{B^c} f(\theta,x)\, d\omega(\theta)}{\int_\Theta f(\theta,x)\, d\omega(\theta)} < 0.5 \\ 0.5 & \text{othervise} \end{cases} \tag{42}$$

is a Bayes measure under the Maximum Uncertainty Principle.

Theorem 6.6. The regular conditional distribution of $\hat{\Theta}$, given $X = x$, $\omega(\cdot \mid x)$, is called the MUP posterior distribution of $\hat{\Theta}$, after obtaining the sample x.

$$\Upsilon_{\hat{\Theta}}(\theta \mid x) = \omega\{\hat{\Theta} \le \theta \mid x\} \tag{43}$$

where $\omega(\cdot \mid x)$ is given by Definition 6.5.

Remark 6.7. Once the MUP posterior density is specified, the posterior mean and variance can be calculated:

$$\mathrm{E}\left[\hat{\Theta}\right] = \int_{\Theta} \theta\, d\Upsilon_{\hat{\Theta}}(\theta \mid x) \tag{44}$$

and

$$\mathrm{V}(\hat{\Theta}) = \mathrm{E}\left[\left(\hat{\Theta} - \mathrm{E}\left[\hat{\Theta}\right]\right)^2\right] = \int_{\Theta} \left(\theta - \mathrm{E}\left[\hat{\Theta}\right]\right)^2 d\Upsilon_{\hat{\Theta}}(\theta \mid x) \tag{45}$$

respectively.

Example 6.8. **:** Let the uncertain prior be

$$\omega(\theta) = \begin{cases} 0 & \text{if } \theta \le a \\ \dfrac{\theta - a}{2(b - a)} & \text{if } a < \theta \le b \\ \dfrac{\theta + c - 2b}{2(c - b)} & \text{if } b < \theta \le c \\ 1 & \text{otherwise} \end{cases} \tag{46}$$

hence

$$\frac{d\omega(\theta)}{d\theta} = \begin{cases} \dfrac{1}{2(b-a)} & \text{if } a < \theta \le b \\ \dfrac{1}{2(c-b)} & \text{if } b < \theta \le c \\ 0 & \text{if } \theta \le a \text{ or } \theta > c \end{cases} \tag{47}$$

The sample of size n is drawn *i.i.d.* from a normal family $N(\theta,\sigma_0^2)$, where the variance σ_0^2 is given. The sample is $x=(x_1,x_2,\cdots,x_n)$. Then

$$\frac{dP_\theta}{dx}=\prod_{i=1}^{n}\phi\left(-\frac{x_i-\theta}{\sigma_0}\right) \tag{48}$$

where

$$\phi(y)=\frac{1}{\sqrt{2\pi}}e^{-\frac{y^2}{2}} \tag{49}$$

Note that

$$\prod_{i=1}^{n}\phi\left(\frac{x_i-\theta}{\sigma_0}\right)=\prod_{i=1}^{n}\frac{1}{\sqrt{2\pi}\sigma_0}\exp\left(-\frac{(x_i-\theta)^2}{2\sigma_0^2}\right) \tag{50}$$

Since

$$\sum_{i=1}^{n}(x_i-\theta)^2=n\left(\theta^2-2\bar{x}_n\theta\right)+Q^2$$
$$=n(\theta-\bar{x}_n)^2+Q^2-\left(\frac{\bar{x}_n}{2}\right)^2$$

where

$$\bar{x}_n=\frac{1}{n}\sum_{i=1}^{n}x_i,\ Q^2=\sum_{i=1}^{n}x_i^2 \tag{51}$$

thus

$$\frac{dP_\theta}{dx}=\left(\frac{1}{\sqrt{2\pi}\sigma_0}\right)^n e^{-(\theta-\bar{x}_n)^2/\left(2\sigma_0^2/n\right)}e^{-\left(Q^2-(\bar{x}_n/2)^2\right)} \tag{52}$$

The absolute distribution for x is

$$\begin{aligned}p(x)&=p^*(x)e^{-\left(Q^2-(\bar{x}_n/2)^2\right)}\\&=\frac{1}{2(b-a)}\left(\Phi\left(\frac{b-\bar{x}_n}{\sigma_0}\right)-\Phi\left(\frac{a-\bar{x}_n}{\sigma_0}\right)\right)\\&+\frac{1}{2(c-b)}\left(\Phi\left(\frac{c-\bar{x}_n}{\sigma_0}\right)-\Phi\left(\frac{b-\bar{x}_n}{\sigma_0}\right)\right)\end{aligned} \tag{53}$$

The posterior density of θ given $x=(x_1,x_2,\cdots,x_n)$ is

$$\gamma(\theta\,|\,x)=\begin{cases}\dfrac{1}{2(b-a)\,p^*(x)}\phi\left(\dfrac{\theta-\bar{x}_n}{\sigma_0/\sqrt{n}}\right) & \text{if } a\le\theta<b\\ \dfrac{1}{2(c-b)\,p^*(x)}\phi\left(\dfrac{\theta-\bar{x}_n}{\sigma_0/\sqrt{n}}\right) & \text{if } b\le\theta<c\\ 0 & \text{otherwise}\end{cases} \tag{54}$$

Example 6.9. Suppose that a random sample of size n is taken from the electronic system lifetime with density

$$\frac{dP_\lambda}{dt}=\prod_{i=1}^{n}\lambda e^{-\lambda t_i}=\lambda^n e^{-\lambda T_n} \tag{55}$$

where

$$T_n=\sum_{i=1}^{n}t_i$$

Let us further assume the uncertain prior density

$$\frac{d\omega(\theta)}{d\theta}=\begin{cases}\dfrac{1}{2(b-a)} & \text{if } a<\theta\le b\\ \dfrac{1}{2(c-b)} & \text{if } b<\theta\le c\\ 0 & \text{if } \theta\le a \text{ or } \theta>c\end{cases}$$

Note

$$\begin{aligned}p(t)&=p^*(t)\\&=\frac{1}{2(b-a)}\left(b^n e^{-bT_n}-a^n e^{-aT_n}\right)\\&+\frac{1}{2(c-b)}\left(c^n e^{-cT_n}-b^n e^{-bT_n}\right)\end{aligned} \tag{56}$$

It is easy to obtain that the posterior density is

$$\omega(\lambda\,|\,t)=\begin{cases}\dfrac{1}{2(b-a)\,p^*(t)}\lambda^n e^{-\lambda T_n} & \text{if } a<\theta\le b\\ \dfrac{1}{2(c-b)\,p^*(t)}\lambda^n e^{-\lambda T_n} & \text{if } b<\theta\le c\\ 0 & \text{if } 0\le\theta\le a \text{ or } \theta>c\end{cases} \tag{57}$$

In case of uncertain observations are adequately large, an MUP asymptotic Bayesian analysis can be carried forward immediately by noting the asymptotic normal distribution $N\left(\mathrm{E}\left[\widehat{\Theta}\right],\mathrm{V}\left[\widehat{\Theta}\right]\right)$.

7. Conclusion

In this paper, we review the newly proposed axiomatic uncertain measure theory and further introduce a measure theoretic treatment of uncertainty decision theory. We further explore the characteristics of the uncertainty decision theory. In terms of our investigations, we emphasize the fundamental mechanism of uncertainty distributions and the impacts on the general uncertainty decision processes. In the discrete uncertainty decision example, the characteristic of uncertainty distribution is intrinsic and unique because the probabilistic discrete distributions never have such features. We develop the MUP Bayes formula in continuous case.

Our efforts in this paper reveal that under general uncertainty, the decision may adopt a framework similar to statistical decision theoretic framework. However, the sub- σ -additive characteristic imposes intrinsic and unique features to the uncertainty distribution and its expectation, and thus the general uncertainty decision making is more computational demanding.

References

[1] Carvalho, H. & Machado, V.C. (2006). Fuzzy set theory to establish resilient production systems. *Proc. of IIE Annual Conference and Exhibition*.

[2] DeGroot, M.H. (1970). *Optimal Statistical Decisions.* McGraw-Hill, Inc, New York, USA

[3] Ferguson, T.S. (1967). *Mathematical Statistics. A Decision Theoretic Approach.* Academic Press Inc, New York, USA.

[4] Kolmogorov, A.N. (1950). *Foundations of the Theory of Probability*. Translated by Nathan Morrison, Publisher: Chelsea, New York

[5] Liu, B.D. (2007). *Uncertainty Theory: An Introduction to Its Axiomatic Foundations*. 2nd Edition; Berlin: Springer-Verlag Heidelberg.

[6] Liu, B.D. (2010). *Uncertainty Theory*. 3nd Edition (Drafted version).

[7] Peng Z.X. & Iwamura, K. (2009). A sufficient and necessary condition of uncertainty distribution.http://orsc.edu.cn/online/090305.pdf .

[8] Primas, H. (1999). Basic Elements and Problems of Probability Theory. *Journal of Scientific Exploration*, Vol. 13, No. 4, pp. 579–613.