فصل پنجم

احتمال و عدم قطعیت

Probability and Uncertainty

- Since Zadeh proposed the fuzzy concept in 1965, there have been many discussions about the relationship between the fuzzy theory and probability theory.
- Both theories express uncertainty, have their values in the range of [0, 1], and have similarities in many aspects.
- In this section, we review the definitions of the two theories and compare them.

Probability Theory

Probability theory deals with the probability for an element to occur in universal set. We call the element as event and the set of possible events as sample space. In the sample space, the elements, i.e., events are mutually exclusive.

Example 7.1 When we play a six-side-dice, the sample space is $S=\{1, 2, 3, 4, 5, 6\}$. Among these six events, only one event can occur. The probability for any of these six events is 1/6.

An event might contain multiple elements. Consider two events A and B as follows.

$$A = \{1,3,5\}, \quad B = \{1,2,3\}.$$

The union and the intersection of these two events are,

$$A \cup B = \{1,2,3,5\}, \quad A \cap B = \{1,3\}$$

and the complement of the event A is

$$\overline{A} = \{2,4,6\}. \quad \square$$

Definition (Probability distribution) To express the probability of events to occur in the sample space, we can define the ***probability distribution*** as follows.

The probability distrivution P is a numerically valued function that assigns a number $P(A)$ to event A so that the following axioms hold. In the axioms, S denotes the sample space.

i) $0 \le P(A) \le 1$,

ii) $P(S) = 1$,

iii) For the mutually exclusive events A_1, A_2, … (that is, for any $i \neq j$, $A_i \cap A_j = \Phi$)

$$P(\bigcup_{i=1}^{\infty} A_i) = \sum_{i=1}^{\infty} P(A_i). \quad \square$$

If the probability distribution is defined, the following properties are satisfied.

(1) $P(A \cup B) = P(A) + P(B) - P(A \cap B)$.

(2) $P(A \cup B) = P(A) + P(B)$, if $A \cap B = \varnothing$.

(3) $P(A) + P(\overline{A}) = 1$.

Now assume that there are two sample spaces S and S', and an event A can occur in S and B in S'. When these events can occur in the mutually independent manner, the ***joint probability*** $P(AB)$ for both A and B to occur is

$$P(AB) = P(A) \cdot P(B).$$

The ***conditional probability*** $P(A \mid B)$ for A provided that the event B has occurred is

$$P(A \mid B) = \frac{P(AB)}{P(B)}.$$

7.1.2 Possibility Distribution

Fuzzy set A is defined on an universal set X and each element in the universal set has its membership degree in [0, 1] for the set A.

$$\mu_A(x) > 0 \quad \text{for} \quad x \in A$$
$$= 0 \quad \text{otherwise.}$$

The membership function μ_A can be defined as a possibility distribution function for the set A on the universal set X. The possibility of element x is denoted as $\mu_A(x)$ and these possibilities define the fuzzy set A.

We know the probability distribution P is defined on a sample space S and the sum of these probabilities should be equal to 1. Meanwhile, the possibility distribution is defined on an universal set X but there is not limit for the sum.

مثال (۳-۵) - عبارت «شخصی در صبحانه X تخم مرغ می خورد» را در نظر بگیرید :

$X = \{1, 2, \ldots\}$

یک تابع توزیع امکان و همین طور یک تابع توزیع احتمال برای X قابل ارائه است. تابع توزیع امکان $\pi_x(u)$ می تواند به عنوان «درجهٔ آسان بودن خوردن u تخم مرغ در صبحانه» تعبیر شود. در حالی که توزیع احتمال زمانی به دست می آید که مشاهدات مربوط به خوردن تخم مرغ توسط آن شخص مثلاً برای ۱۰۰ روز جمع آوری و فراوانی ها محاسبه شوند. مقادیر $\pi_x(u)$ و $P_x(u)$ می توانند به صورت ذیل نمایش داده شوند :

u	۱	۲	۳	۴	۵	۶	۷	۸
$\pi_x(u)$	۱	۱	۱	۱	۰.۸	۰.۶	۰.۴	۰.۲
$P_x(u)$	۰.۱	۰.۸	۰.۱	۰	۰	۰	۰	۰

Example 7.2 Suppose the following proposition :

"Sophie has x sisters."

$x \in N = \{1, 2, 3, 4, \ldots 10\}$.

Both probability distribution and possibility distribution can be used to define the variable x in N (Table 7.1). If we use the probability distribution P, the probability of having x sister(s) is defined by $P(x)$. By the possibility distribution μ, we define the possibility with x sisters as $\mu_A(x)$. The set N is considered as a sample space in the probability distribution and as a universal set in the possibility distribution.

Table 7.1. Possibility and Probability

x	1	2	3	4	5	6	7	8	9	10
P(x)	0.4	0.3	0.2	0.1	0	0	0	0	0	0
μ (x)	0.9	1.0	1.0	0.7	0.5	0.2	0.1	0	0	0

We see that the sum of the probabilitis is equel to 1 but that of the possibilities is greater than 1. In (Table 7.1.),

we can see that higher possibility does not always means higher probability.

But lower possibility leads to lower probability.

So we can say that the possibility is the upper bound of the probability.

7.1.3 Comparison of Probability and Possibility

Probability and possibility have something in common: they both describe uncertainty. The possibility can be regarded as the upper bound of probability value. That is, the possibility $\mu(A)$ and probability $P(A)$ of an event A have the following relation.

$$\mu(A) \geq P(A).$$

If the events $A_1, A_2, \ldots, A_n$ are mutually exclusive, the probability of union of these events is equivalent to the sum of the probabilities of each event, and that of intersection is equivalent to the multiplication.

$$P(\cup_i A_i) = \sum_i P(A_i)$$

$$P(\cap_i A_i) = P(A_1) \cdot P(A_2) \cdot \ldots \cdot P(A_n).$$

The possibility for union of those events has the maximum value and that for intersection has the minimum. (Table 7.2) compares the characteristics of possibility with those of probability.

$$\mu(\cup_i A_i) = \operatorname{Max}_i \mu(A_i)$$

$$\mu(\cap_i A_i) = \operatorname{Min}_i \mu(A_i).$$

Table 7.2. Comparison of Possibility and Probability

	Possibility	Probability
Domain	Universal set X	Sample space S
Range	[0,1]	[0,1]
Constraints	none	$\sum_i P(A_i)=1$
Union	$\mu(\bigcup_i A_i)=\operatorname{Max}_i \mu(A_i)$	$P(\bigcup_i A_i)=\sum_i P(A_i)$
Intersection	$\mu(\bigcap_i A_i)=\operatorname{Min}_i \mu(A_i)$	$P(\bigcap_i A_i)=P(A_1)\cdot P(A_2)\cdot\ldots\cdot P(A_n)$

نتیجه گیری کلی

به عنوان نتیجه گیری کلی از بحث می توان چنین گفت که نظریه مجموعه های فازی، نظریه امکان و نظریه احتمالات جانشین یکدیگر نیستند اما مکمل یکدیگرند. این طور به نظر می رسد که نظریه مجموعه های فازی در زمینه های مختلف نسبت به احتمالات سازگارتر است. البته در هر زمینه ای که لازم است با مجموعه های فازی تحلیل گردد می بایست نظریه مجموعه های فازی با ابزار مدل سازی همراه شود.

۳-۵- اندازه امکان و اندازه الزام

تعریف (۳-۵)- اندازه امکان در مجموعه های کلاسیک : پروفسور زاده مفهوم اندازه امکـان[۷] را به صورت ذیل تعریف کرد :

اندازه امکان تابعی است ($\Pi : A \to [0,1]$) که مجموعه A از فضای جهانی X را به بازه پیوسـته $[0,1]$ تصویر می کند و دارای خصوصیات ذیل است :

الف- $\Pi(0) = 0 \quad , \quad \Pi(X) = 1$

ب- $A \subseteq B \quad \to \quad \Pi(A) \leq \Pi(B)$

ج- $\Pi(\bigcup_{i \in I} A_i) = \sup_{i \in I} \Pi(A_i)$

اندازه امکان می تواند به صورت یگانه با یک تابع توزیـع امکـان ($\pi_x : X \to [0,1]$) و بـه شـرح ذیل تعیین شود.

$$\Pi(A) = \sup_{x \in A} \pi_x \, , \qquad A \subset X$$

و همین طور برای هر عنصر x داریم :

$$\pi_x = \Pi(\{x\}), \quad \forall x \in X$$

مثال (۴-۵)- فرض کنید مجموعه جهانی X مجموعه اعداد صحیح غیر منفی کوچکتر از ۱۱ تعریف شود که به صورت ذیل خواهد بود :

$X = \{0, 1, \ldots, 10\}$

تابع توزیع امکان این که x عددی نزدیک به عدد ۸ باشد(π_x) می تواند به صورت ذیل تعریف شود :

x	۱	۲	۳	۴	۵	۶	۷	۸	۹	۱۰
π_x	۰	۰	۰	۰	۰.۱	۰.۵	۰.۸	۱	۰.۸	۰.۵

حال فرض کنید مجموعه A به صورت ذیل تعریف شود آن گاه امکان وجود یک عدد نزدیک به ۸ در مجموعه A به صورت ذیل محاسبه می شود :

$A = \{2, 5, 9\}$

$\Pi(A) = \sup\{\pi_2, \pi_5, \pi_9\}$

$\Pi(A) = \sup\{0, 0.1, 0.8\}$

$\Pi(A) = 0.8$

خواص ذیل برای اندازه امکان در حالات حدی وجود دارد :

$$\Pi(A \bigcup B) = \max\big(\Pi(A), \Pi(B)\big)$$

$$\Pi(A \bigcap B) = \min\big(\Pi(A), \Pi(B)\big)$$

اگر $\overline{A}$، مکمل مجموعه A در U فرض شود آن گاه:

$$\Pi(A \cup \overline{A}) = \max\big(\Pi(A), \Pi(\overline{A})\big) = 1$$

که بیانگر این است که امکان وقوع A یا $\overline{A}$ قطعی است.

در نظریه امکان، اندازه دیگری نیز قابل تعریف است که از رابطـه عطـف اسـتفاده مـی کنـد و مزدوج اندازه امکان است :

$$N(A \cap B) = \min\big(N(A), N(B)\big)$$

N اندازهٔ الزام[8] نامیده می شود. عبارت ($N(A)=1$) بیـانگر ایـن اسـت کـه A الزامـاً درسـت است. ارتباط مزدوج بین اندازه امکان و اندازه الزام طبق رابطهٔ ذیل برقرار است :

$$\Pi(A) = 1 - N(\overline{A}) \ , \qquad \forall A \subseteq U$$

همچنین اندازهٔ الزام شرط ذیل را تامین می کند :

$$\min\big(N(A), N(\overline{A})\big) = 0$$

ارتباط بین اندازه امکان و اندازه الزام روابط ذیل را ارضاء می کند :

$$\Pi(A) \geq N(A) \ , \qquad \forall A \subseteq U$$

$$N(A) > 0 \ \Rightarrow \ \Pi(A) = 1$$

$$\Pi(A) < 1 \ \Rightarrow \ N(A) = 0$$

مجموعه جهانی U همیشه متناهی فرض می شود.

با داشتن تابع توزیع امکان، اندازه امکان برای مجموعه A از مجموعه جهـانی U طبـق رابطـه ذیل محاسبه می شود :

$$N(A) = \min_{i \in \overline{A}} \{1 - \pi_i\}$$

مثال (۵-۵)- فرض کنید بر اساس تجارب قبلی، تابع توزیع امکان برای اخذ نمره A تا E برای شش دانشجو طبق جدول (۱-۵) تعریف شود.

جدول (۱-۵)- تابع توزیع امکان نمرات برای دانشجویان

	نمره				
دانشجو	A	B	C	D	E
1	0.8	1	0.7	0	0
2	1	0.8	0.6	0.1	0
3	0.6	0.7	0.9	0.1	0
4	0	0.8	0.9	0.5	0
5	0	0	0.3	1	0.2
6	0.3	1	0.3	0	0

سئولات مختلفی با توجه به تابع توزیع امکان تعریف شده قابل پاسخگویی است که در ذیل چند سئوال مطرح و پاسخ ارایه شده است.

سؤال الف : عبارت «دانشجوی ۱ در امتحان بعدی نمره B اخذ خواهد کرد» چقدر قابل اطمینان است؟

در این مورد مجموعه های مورد نظر عبارتند از : $A = \{B\} \quad , \quad \overline{A} = \{A, C, D, E\}$

لذا خواهیم داشت :

$$\Pi(A) = 1$$
$$N(A) = \min\{1 - \Pi_i\} = \min\{\ 0.2, 0.3, 1, 1\} = 0.2$$

در نتیجه امکان این که دانشجوی ۱ در امتحان بعدی نمره B بگیرد ۱ است در حالی که ضرورت یا الزام این که نمره B اخذ کند ۰،۲ است.

سؤال ب : اگر بخواهیم درستی عبارت «یکی از دانشجویان ۱ یا ۲ نمره A یا B خواهند گرفت» را بررسی کنیم، مجموعه U تغییر کرده و شامل دو سطر اول از جدول است که نتایج به صورت ذیل به دست می آید :

$\Pi(A) = \Pi$ (دانشجوی ۱ نمره A یا B یا دانشجوی ۲ نمره A یا B) = 1

$N(A) = 0.3$

سؤال ج : معتبر بودن این که «دانشجوی ۱ نمره C خواهد گرفت» به صورت ذیل بررسی می شود :

$$\Pi(A) = 0.7$$

$$N(A) = 0$$

تعریف (۴-۵)- اندازه امکان در مجموعه های فازی : فرض کنید $\tilde{A}$ یک مجموعه فازی در مجموعه جهانی U باشد و همین طور فرض کنید π_x توزیع امکان مقادیر x در مجموعه جهانی U باشد. آن گاه اندازه امکان مجموعه فازی $\tilde{A}$، ($(\pi_x(\tilde{A})$) به صورت ذیل تعریف می شود :

$$\text{POSS } \{X \text{ is } \tilde{A}\} \hat{=} \pi(\tilde{A})$$

$$\hat{=} \sup_{u \in U} \min\{\mu_{\tilde{A}}(u),\ \pi_x(u)\}$$

مثال (۶-۵) - مثال (۲-۵) که در آن توزیع امکان گزاره «x یک عدد صحیح کوچک است» به صورت ذیل به دست آمد را در نظر بگیرید :

U	۱	۲	۳	۴	۵	۶
π_x	۱	۱	۰.۸	۰.۶	۰.۴	۰.۲

برای مجموعه کلاسیک $A=\{3, 4, 5\}$ اندازه امکان A (امکان این که یک عدد کوچک در A باشد) به صورت ذیل محاسبه می شود :

$$\Pi(A) = \max(0.8, 0.6, 0.4) = 0.8$$

حال اگر $\tilde{A}$ یک مجموعه فازی و به صورت " اعداد صحیحی که کوچک نیستند" تعریف شود :

$$\tilde{A} = \{(3, 0.2), (4, 0.4), (5, 0.6), (6, 0.8), (7, 1), \ldots\}$$

اندازه امکان این که یک عدد صحیح کوچک در $\tilde{A}$ باشد به صورت ذیل محاسبه می شود :

$$\pi(\tilde{A}) = \sup\{\ \min(0.2, 0.8), \min(0.4, 0.6), \min(0.6, 0.4), \min(0.8, 0.2)\ \}$$
$$= \sup\{\ 0.2, 0.4, 0.4, 0.2\ \} = 0.4$$

Fuzzy Event

When dealing with the ordinary probability theory, an event has its precise boundary. For instance, if an event is $A=\{1,3,5\}$, its boundary is sharp and thus it can be represented as a crisp set. When we deal an event whose boundary is not sharp, it can be considered as a fuzzy set, that is, a fuzzy event. For example,

$$B = \text{“ small integer ”} = \{(1, 0.9), (2, 0.5), (3, 0.3)\}$$

How would we deal with the probability of such fuzzy events?

We can identify the probability in two manners:

- One is dealing with the probability as a crisp value(crisp probability)
- and the other as a fuzzy set(fuzzy probability).

7.2.1 Crisp Probability of Fuzzy Event

Let a ***crisp event*** **A** be defined in the space $\Re^n$ All events in the space $\Re^n$ are mutually exclusive and the probability of each event is,

$$P(A) = \int_A dP$$

and for discrete event in the space $\Re^n$,

$$P(A) = \sum_{x \in A} P(x).$$

Let $\mu_A(x)$ be the ***membership function*** of the event(set) A and the ***expectation*** of $\mu_A(x)$ be $E_P(\mu_A)$. Then the following relationship is satisfied.

$$P(A) = \int_A \mu_A dP = E_P(\mu_A)$$

for discrete elements,

$$P(A) = \sum_{x \in A} \mu_A(x) P(x).$$

Definition (Crisp probability of fuzzy event) Let event A be a ***fuzzy event*** or a fuzzy set considered in the space $\Re^n$:

$$A = \{(x, \mu_A(x)) \mid x \in \Re^n\}$$

The probability for this fuzzy set is defined as follows:

$$P(A) = \int_A \mu_A dP = E_P(\mu_A)$$

and alternatively,

$$P(A) = \sum_{x \in A} \mu_A(x) P(x) . \quad \square$$

Example 7.3 Assume that the sample space $S=\{a,b,c,d\}$ is given as in (Fig 7.1.) Each element is mutually exclusive, and each probability is given as,

$P(a) = 0.2$, $P(b) = 0.5$, $P(c) = 0.2$, $P(d) = 0.1$.

Think about a crisp event $A=\{a,b,c\}$ in the sample space S with its characteristic function given as (Fig 7.2),

$$\mu_A(a) = \mu_A(b) = \mu_A(c) = 1, \quad \mu_A(d) = 0$$

the probability of the crisp event A can be calculated from the following procedure.

$$P(A) = 1 \cdot 0.2 + 1 \cdot 0.5 + 1 \cdot 0.2 = 0.9 .$$

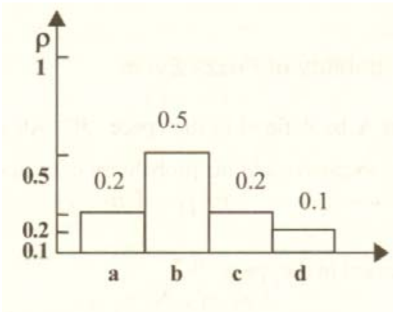


Fig. 7.1. Sample space S and Probability

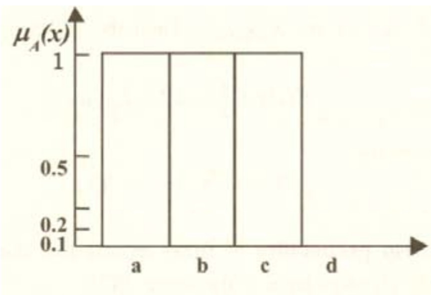


Fig. 7.2. Crisp event A

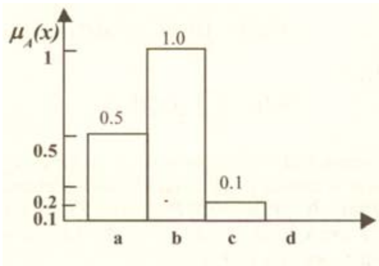


Fig. 7.3. Fuzzy event A

now consider the event A as a fuzzy event (Fig 7.3). That is,

$$A = \{(a,0.5),(b,1),(c,0.1)\}.$$

The crisp probability for the fuzzy event A is,

$$P(A) = 0.5 \cdot 0.2 + 1 \cdot 0.5 + 0.1 \cdot 0.2 = 0.62.$$

7.2.2 Fuzzy Probability of Fuzzy Event

Given the following fuzzy event in the sample space S,

$$A = \{(x, \mu_A(x)) \mid x \in S\}.$$

The α-cut set of the event(set) A is given as the following crisp set.

$$A_\alpha = \{x \mid \mu_A(x) \geq \alpha\}.$$

The probability of the α-cut event is as following:

$$P(A_\alpha) = \sum_{x \in A_\alpha} P(x).$$

Here, A_α is the union of mutually exclusive events. The probability of A_α is the sum of the probability of each event in the α-cut set A_α. For the probability of the α-cut event, we can say that

"The possibility of the probability of set A_α to be $P(A_\alpha)$ is α"

taking this interpretation, there are multiple cases for the fuzzy probability $P(A)$ according to the value α.

Definition (Fuzzy probability of fuzzy event) Fuzzy event A, its α-cut event A_α and the probability $P(A_\alpha)$ are provided from the above procedure. The fuzzy probability P(A) is defined as follows:

$$P(A) = \{(P(A_\alpha), \alpha) \mid \alpha \in [0,1]\}. \quad \square$$

Of course, the value of α is an element in the level set of fuzzy set *A*. We used the same kind of interpretation when we discussed about the fuzzy cardinality of fuzzy set in chapter 1.

Example 7.4 Assume the probability of each element in the sample space *S*={a,b,c,d} as shown in (Fig 7.4.)

$$P(a) = 0.2, P(b) = 0.3, P(c) = 0.4, P(d) = 0.1$$

A fuzzy event *A* is given in (Fig 7.5.)

$$A = \{(a,1), (b,0.8), (c,0.5), (d,0.3)\}$$

taking the α-cut event A_α , we get crisp events.

$$A_{0.3} = \{a,b,c,d\}$$

$$A_{0.5} = \{a,b,c\}$$

$$A_{0.8} = \{a,b\}.$$

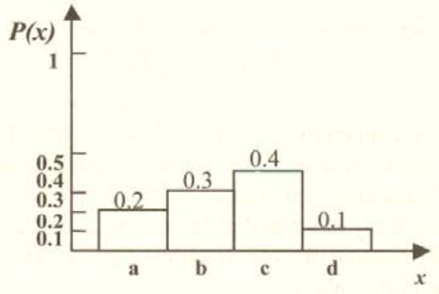


Fig. 7.4. Sample space $S = \{a,b,c,d\}$

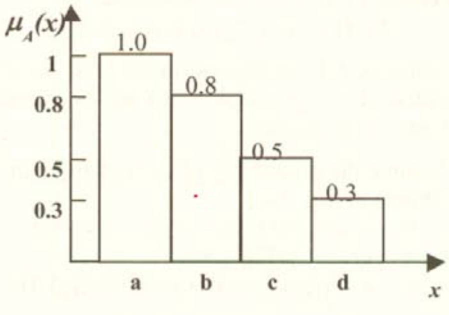


Fig. 7.5. Fuzzy event *A*

$$A_1 = \{a\}$$

since these are crisp events, we can easily calculate the probabilities of each α-cut event.

$$P(A_{0.3}) = 0.2 + 0.3 + 0.4 + 0.1 = 1$$
$$P(A_{0.5}) = 0.2 + 0.3 + 0.4 = 0.9$$
$$P(A_{0.8}) = 0.2 + 0.3 = 0.5$$
$$P(A_1) = 0.2 .$$

Now, the possibility for this fuzzy event A to be $P(A_\alpha)$ is α, and the probability of the fuzzy event A is given as follows (Fig7.6):

$$P(A) = \{(1,0.3),(0.9,0.5),(0.5,0.8),(0.2,1)\} . \quad \square$$

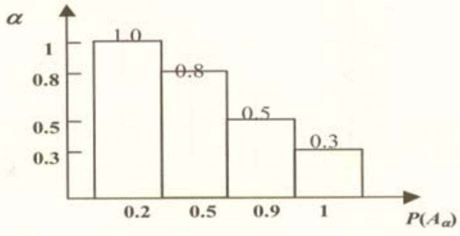


Fig. 7.6. Fuzzy probability

Uncertainty

7.3.1 Uncertainty Level of Element

Suppose students a, b, c and d take the entrance examination for college A. The possibility for each student to enter the college is 0.2, 0.5, 0.9 and 1 respectively. At this point, the possibilities can be identified as a fuzzy set. That is, let the college be a fuzzy set A and a, b, c and d be the elements of A. The possibilities to be contained in A can be expressed as the values of membership function (Fig 7.7).

$$\mu_A(a) = 0.2 ,\quad \mu_A(b) = 0.5 ,\quad \mu_A(c) = 0.9 ,\quad \mu_A(d) = 1 .$$

When discussing the possibilities to be in A, *which of these elements a, b, c and d has the largest uncertainty?*

First, the element d has the concrete possibility, we say it has the least uncertainty. The element a, on the other hand, has almost no possibility to pass the examination. This student doesn't expect too much for his success and we might say he has relatively less uncertainty. However, the student b might have the most uncertainty since he has the possibility 0.5.

- When the possibility is near to 0.5, the uncertainty gets the highest.
- As the possibility approaches 0 or 1, the uncertainty decreases.

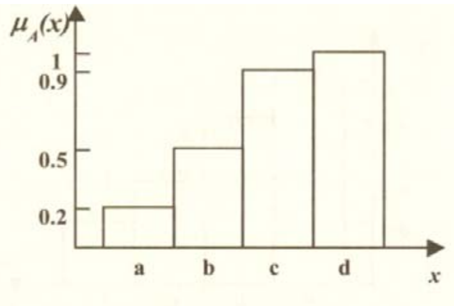


Fig. 7.7. Fuzzy set A

7.3.2 Fuzziness of Fuzzy Set

Suppose students a, b, c and d apply to B college and consider B as a fuzzy set, the possibilities for those students to succeed in the entrance examination are shown in the following membership functions (Fig 7.8).

$$\mu_B(a) = 0.4, \quad \mu_B(b) = 0.7, \quad \mu_B(c) = 0.6, \quad \mu_B(d) = 0.5.$$

Comparing the fuzzy sets A and B, which one is more uncertain?

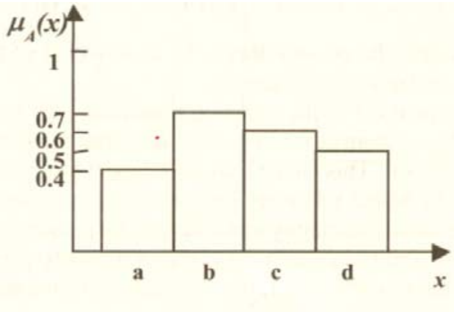


Fig. 7.8. Fuzzy set B

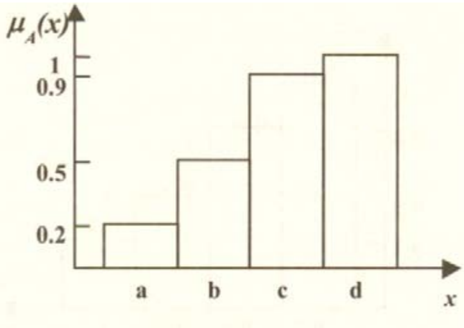


Fig. 7.7. Fuzzy set A

➢ Comparisons element by element, the elements of *B are more uncertain* (the values of membership functions are closer to 0.5).

➢ So, the fuzzy set *B* has more uncertain states comparing with *A*.

➢ *So as to speak, B is relatively* more fuzzy.

➢ If we consider another fuzzy set C, each element of which having its membership degree 0.5, the fuzzy set C has the largest degree of fuzziness (uncertainty).

Measure of Fuzziness

Fuzziness is termed for the case when representation of uncertainty level is needed. The function for this fuzziness is called ***measure of fuzziness***. The function f denoting the measure of fuzziness is

$$f : P(X) \rightarrow \mathrm{R}.$$

In the function, $P(X)$ is the power set gathering all subsets of the universal set X, and R is the real number domain. The function f grants the real value $f(A)$ to the subset A of X, and the value indicates the fuzziness of set A, there are three conditions for the measures to observe.

(Axiom F1) $f(A) = 0$ iff A is a crisp set.

The fuzziness should have the value 0 when the set is crisp (Fig 7.9).

(Axiom F2) When the uncertainty of A is less than that of B , the measured value $f(A)$ should be less than $f(B)$. When we denote it by $A < B$ and say that A is sharper than B, the following relation should be satisfied.

$$f(A) \leq f(B).$$

The relation implies the ***monotonicity*** property.

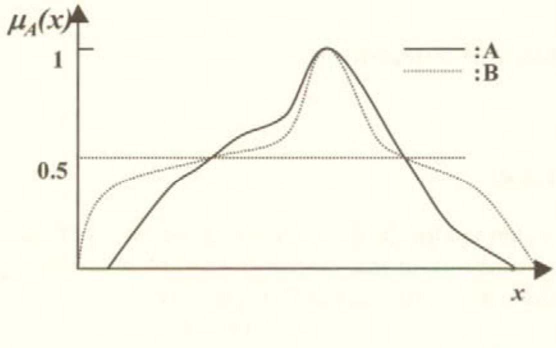


Fig. 7.10. When A is sharper than B, $f(A) \leq f(B)$.

Example 7.5 When two fuzzy sets A and B are presented (Fig 7.10), the fact that A is sharper than B is defined as following. That is, if $A < B$ holds, then the followings are satisfied.

i) when $\mu(x) \leq \frac{1}{2}$, $\mu_A(x) \leq \mu_B(x)$

ii) when $\mu(x) \geq \frac{1}{2}$, $\mu_A(x) \geq \mu_B(x)$

By the axiom F2, the relation $f(A) \leq f(B)$ should be satisfied. □

(Axiom F3) If the fuzziness is the maximum, the measure $f(B)$ should have the maximum value.

Example 7.6 For the deep understanding of the axiom F3, think of the case that the uncertainty is the maximum. If the membership degree of each element x in fuzzy set A is 0.5, the uncertainty has the maximum. That is, for all elements $x \in A$, $\mu_A(x)=0.5$, and then the fuzziness measure $f(A)$ is also the maximum (Fig 7.11). □

Fig. 7.11. Maximum uncertainty $f(A)$.

7.4.2 Measure using Entropy

In this section, a fuzziness measure based on Shannon's entropy is explained. Shannon's entropy is widely used in measuring the amount of uncertainty or information, and is considered as the fundamental theory in the information theory.

Definition (Shannon's entropy)

$$H(P(x)) = -\sum_{x \in X} P(x)\log_2 P(x), \quad \forall x \in X. \quad \square$$

In the rest of the section, $P(x)$ denotes the probability distribution in the universal set X for all $x \in X$.

Example 7.7 To know more about the entropy, consider the following examples with two probability distributions P and P'.

i) Probability distribution P

For the universal set $X=\{a,b,c\}$, the probability distribution P is given as,

$$P(a)=1/3,\ P(b)=1/3,\ P(c)=1/3,$$

$$P(a)+P(b)+P(c)=1.$$

The probabilities of all elements in X are equal to each other, and the sum of the probabilities is 1. The uncertainty of one element's occurrence is measured by the Shannon's entropy.

$$H(P)=-\left[\tfrac{1}{3}\log_2\tfrac{1}{3}+\tfrac{1}{3}\log_2\tfrac{1}{3}+\tfrac{1}{3}\log_2\tfrac{1}{3}\right]$$

$$=-\log_2\tfrac{1}{3}=\log_2 3=1.6.$$

ii) Probability distribution P'

Assume there is a probability distribution P' for X as follows.

$$P'(a)=1/2,\ P'(b)=1/2,\ P'(c)=0$$

The uncertainty is,

$$H(P')=-\left[\tfrac{1}{2}\log_2\tfrac{1}{2}+\tfrac{1}{2}\log_2\tfrac{1}{2}+0\right]$$

$$=-\log_2\tfrac{1}{2}=\log_2 2=1.$$

The uncertainty for P is greater than that of P' (Fig 7.12). □

In the above example, when the probability distribution is P', the possibility for the element c to occur is 0. So we do not need to consider c. But in the probability distribution P, we need to consider the three elements, a,b,c. So the uncertainty $H(P)$ is greater than $H(P')$. When there are only two events and the probability of each event is 0.5, the amount of information is the maximum at 1 as in the case P'

Since the Shannon's entropy is based on the probability distribution, the total probabilities of all elements is 1.

$$\sum_X P(x) = 1.$$

But for fuzzy sets, this restriction is unnecessary. If a fuzzy set A is defined in the universal set X by a membership function $\mu_A(x)$, the following restriction is not required.

$$\sum_X \mu_A(x) = 1 \quad \text{(not necessary).}$$

Definition (measure of fuzziness)

$$f(A) = -\sum_{x \in X} \left[\mu_A(x) \log_2 \mu_A(x) + (1 - \mu_A(x)) \log_2 (1 - \mu_A(x)) \right].$$

This is the sum of the uncertainties of a fuzzy set A defined by the membership function $\mu_A(x)$ and its complement $\overline{A}$ defined by$[1-\mu_A(x)]$

The normalized measure $\hat{f}(A)$ of the measure $f(A)$ is defined as the following

$$\hat{f}(A) = \frac{f(A)}{|X|}.$$

In the above,$|X|$ denotes the cardinality of the universal set X and the normalized measure observes this relation

$$0 \le \hat{f}(A) \le 1$$

this measure satisfies the axiom (F1) and (F2).

Example 7.8 Suppose that there are two fuzzy sets A and A' in $X=\{a,b,c\}$ (Fig 7.13).

i) Assume that a fuzzy set A is given as

$$A = \{(a,0.5),(b,0.2),(c,1)\}.$$

The fuzziness of the fuzzy set A is,

$$f(A) = -(0.5\log_2 0.5 + 0.5\log_2 0.5 + 0.2\log_2 0.2 + 0.8\log_2 0.8 + 1\log_2 1 + 0)$$

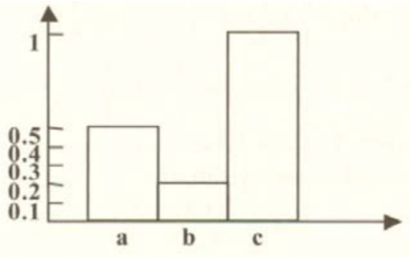
$$= -(\log_2 \frac{1}{2} + \frac{1}{5}\log_2 \frac{1}{5} + \frac{4}{5}\log_2 \frac{4}{5})$$

$$= \log_2 2 + \frac{1}{5}\log_2 5 + \frac{4}{5}\log_2 \frac{5}{4}$$

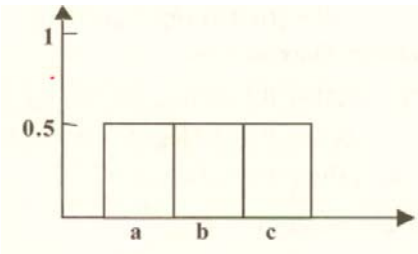
$$= \log_2 5 - 0.6 = 1.7\,.$$

and the normalized measure yields,

$$\hat{f}(A) = \frac{f(A)}{|X|} = \frac{1.7}{3} = 0.57\,.$$



fuzzy set A



fuzzy set A'

ii) The fuzzy set A' is given as follows.

$$A' = \{(a,0.5),(b,0.5),(c,0.5)\}$$

The fuzziness for the fuzzy set A' is,

$$f(A') = -(0.5\log_2 0.5 + 0.5\log_2 0.5 + 0.5\log_2 0.5 + 0.5\log_2 0.5 + 0.5\log_2 0.5 + 0.5\log_2 0.5)$$
$$= -(3\log_2 0.5) = 3\log_2 2 = 3,$$

and the normalized measure is,

$$\hat{f}(A) = \frac{f(A')}{|X|} = \frac{3}{3} = 1.$$

The membership degrees of all elements in A' are 0.5. So the uncertainty of the fuzzy set A' is larger. Consequently, the fuzziness of A' is greater than that of A

$$f(A) < f(A'). \quad \square$$

We state that the uncertainty is the largest when the membership degrees are all 0.5. And the normalized fuzziness of such fuzzy set is 1.

7.4.3 Measure using Metric Distance

Another measure of fuzziness is the one that is based on the concept of metric distance. We talked about Hamming distance and Euclidean distance in sec 2.5. A crisp set C that corresponds to a fuzzy set A is introduced for the distance measure (Fig 7.14).

$$\mu_C(x) = 0 \quad \text{if } \mu_A(x) \le \frac{1}{2}$$
$$\mu_C(x) = 1 \quad \text{if } \mu_A(x) > \frac{1}{2}$$

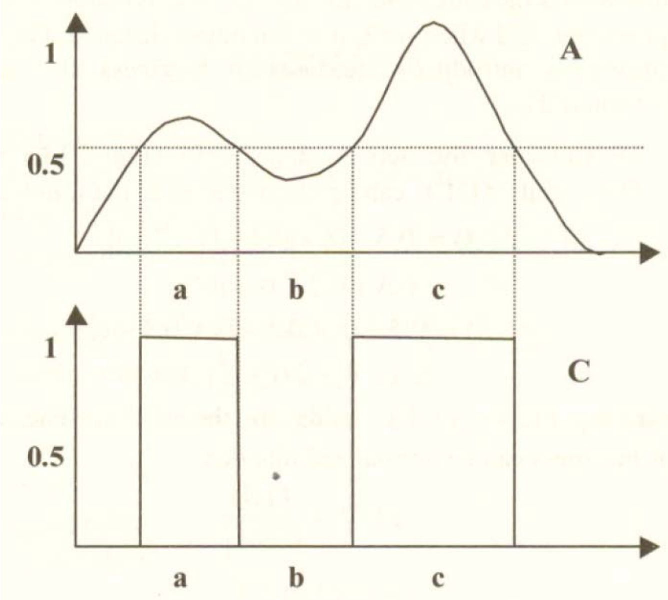


Fig. 7.14. Fuzzy set A and its corresponding Crisp set C

Definition (Hamming distance) The measure of fuzziness $f(A)$ is expressed as,

$$f(A) = \sum_{x \in X} |\mu_A(x) - \mu_C(x)|. \qquad \square$$

Definition (Euclidean distance) The measure of fuzziness $f(A)$ is,

$$f(A) = \left(\sum_{x \in X} [\mu_A(x) - \mu_C(x)]^2 \right)^{\frac{1}{2}}. \qquad \square$$

The closer to 0.5 the values of membership function are, the larger the fuzziness is measured.

Definition (Minkowski distance) Generalizing Hamming distance and Euclidean distance, the following ***Minkowski's measure*** yields

$$f_w(A) = \left(\sum_{x \in X} |\mu_A(x) - \mu_C(x)|^w \right)^{\frac{1}{w}}. \qquad \square$$

The Minkowski's measure holds for $w \in [1, \infty]$. When w=1, it becomes Hamming distance, and when w=2, it is Euclidean distance. We easily see that the previously introduced measures of fuzziness also satisfy the axioms (F1) and (F2).

Example 7.9 Consider the sets A and A' in (Fig 7.13.) Hamming distances $f(A)$ and $f(A')$ can be calculated as in the following.

$$f(A) = |0.5-0| + |0.2-0| + |1-1|$$
$$= 0.5 + 0.2 + 0 = 0.7$$
$$f(A') = |0.5-0| + |0.5-0| + |0.5-0|$$
$$= 0.5 + 0.5 + 0.5 = 1.5 .$$

The relation $f(A) < f(A')$ holds in the above. The Hamming measure of fuzziness can be normalized like this.

$$\hat{f}(A) = \frac{f(A)}{0.5|X|}$$
$$0 \le \hat{f}(A) \le 1 .$$

The symbol $|X|$ denotes the cardinality of the universal set X. The normalized measure of the previous example is

$$\hat{f}(A) = \frac{0.7}{0.5 \times 3} = 0.47 . \qquad \square$$