A Preliminary Study of the Theoretical Basis of the Fuzzy Set

Author : Zhang Nanlun

Address : Dept. of Automation,
Wuhan Institute of Building Materials,
Wuhan, Hubei,
The people's Republic of China.

SUMMARY

1. Introduction

The basis of fuzzy set theory is the concept of the membership function. It would appear desirable to consider the following questions : what is the membership function? what is the true meaning of the membership function? How can membership functions be found? what is the relationship between fuzzy set theory and classical probability theory?

These questions have already been discussed. After having solved the questions, it will be possible to construct the stable basis of fuzzy set theory.

2. The membership and probability characteristics of random appearances

Many random appearances with the same properties can be studied by considering the probability and membership characteristics.

The random appearances obey certain laws of probability distribution which are refered to as the probability characteristics of random appearances.

Similarly, the random appearances obey certain law of membership distribution which are referred to as the membership characteristics of random appearances.

By studying the probability characteristics of random appearances, the statistical laws of random appearances may be found.

By studing the membership characteristicsof random appearances, the new explanations of random appearances may be given.

For the convenience of explanation, itbe considered for the moment that the fuzzy appearance may fall into the category of random appearances. The reason for this will be explained in the 7th section.

3. Finding membership functions by a statistical method

The membership function may be obtained by the original statistical method. However, a new explanation must be given for this method.

Example 3-1. After 129 students of Wuhan Institute of Building Materials thought over the implications of the words " young people ", the limits given by each were as follows (the unit is age) :

18-25	17-30	17-28	18-25	16-35	14-25
18-30	18-35	18-35	16-25	15-30	18-35
17-30	18-25	18-35	20-30	18-30	16-30
20-35	18-30	18-25	18-35	15-25	18-30
15-28	16-28	18-30	18-30	16-30	18-35
18-25	18-30	16-28	18-30	16-30	16-28
18-35	18-35	17-27	16-28	15-28	18-25
19-2[illegible]	15-30	15-26	17-25	15-36	18-30
17-30	18-35	16-35	16-30	15-25	18-28
16-30	15-28	18-35	18-30	17-28	18-35
15-28	15-25	15-25	15-25	18-30	16-24
15-25	16-32	15-27	18-35	16-25	18-30
16-28	18-30	18-35	18-30	18-30	17-30
18-30	18-35	16-30	18-28	17-25	15-30
18-25	17-30	14-25	18-26	18-29	18-35
18-28	18-35	18-25	16-35	17-29	18-25
17-30	16-28	18-30	16-28	15-30	18-30
15-30	20-30	20-30	16-25	17-30	15-30
18-30	16-30	18-28	15-35	16-30	15-30
18-35	18-35	18-30	17-30	16-35	17-30
15-25	18-35	15-30	15-25	15-30	18-30
17-25	18-29	18-28			

In accordance with normal statistical definitions:

Frequency: Numbers of the sample falling into each class boundary.

Relative frequency: The ratio of frequency to the sample size.

Sample size = 129

The frequency distribution obtained is as follows:

Orders	Class Boundaries	Frequency	Relative frequency
1	13.5-14.5	2	0.0155
2	14.5-15.5	27	0.2093
3	15.5-16.5	51	0.3953
4	16.5-17.5	67	0.5194
5	17.5-18.5	124	0.9612
6	18.5-19.5	125	0.9690
7	19.5-20.5	129	1.0
8	20.5-21.5	129	1.0
9	21.5-22.5	129	1.0
10	22.5-23.5	129	1.0
11	23.5-24.5	129	1.0
12	24.5-25.5	128	0.9922
13	25.5-26.5	103	0.7984
14	26.5-27.5	101	0.7829
15	27.5-28.5	99	0.7674
16	28.5-29.5	80	0.6202
17	29.5-30.5	77	0.5969
18	30.5-31.5	27	0.2093
19	31.5-32.5	27	0.2093
20	32,5-33.5	26	0.2016
21	33.5-34.5	26	0.2016
22	34.5-35.5	26	0.2016
23	35.5-36.5	1	0.0078
	Σ		13.6589

The relative frequency histogram is shown in figure 3-1.

It may again be explained why "the relative frequency " can show the possibility of membership degree as well as the probability under certain conditions. In fact the " relative frequency " is shown as the estimate of the degree of generalized possibility. Thus, this concept can represent the meanings of the probability and the possibility of membership degree.

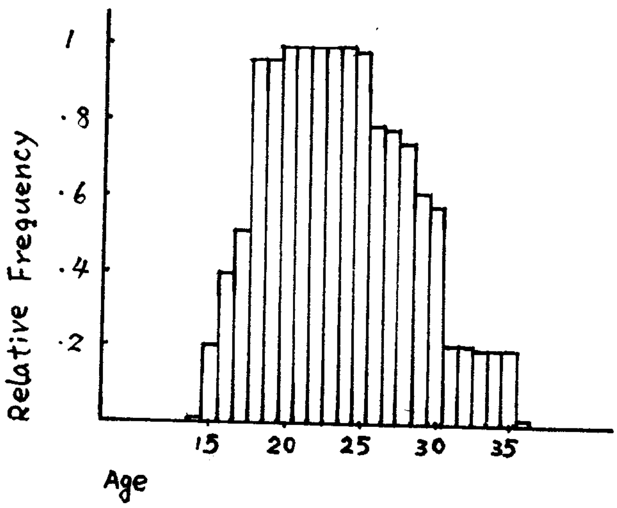


Figure 3-1. Statistical distribution curve of the concept " young people"

4. The logical basis of this kind of statistical method

It would appear appropriate at this stage to recall the logical basis of statistical methods in probability theory.

In probabilistic statistics, the observations can be handled by constructing a frequency table which classifies the set of observations according to the number of variables falling within certain limits (class boundaries).

For each outcome its characteristic function is as follows:

$$A(\lambda) = \begin{cases} 1 & (I) \\ 0 & (1, 2, \ldots I-1, I+1, \ldots N). \end{cases}$$

Where, N= number of class boundaries,

$1 \leq I \leq N.$

That is to say, if some outcome is within class boundary I, then it is not within class boundaries 1 or 2......or I-1 or I+1......or N.

This kind of statistical method is constructed on the basis of two-valued logic of the real point.

Fuzzy statistics uses the same methods as probabilistic statistics. However each outcome may be shown with an interval.

The characteristic function of each outcome may be shown as follows:

$$A(\lambda)=\begin{cases} 1 & (I, I+1, \dots J) \\ 0 & (1,2,\dots I-1, J+1,\dots N) \end{cases}$$

where $I < J$,

$1 \leqslant I, \quad J \leqslant N.$

That is to say, If some outcome is within class boundaries I and I+1...and J, then it cannot be within class boundaries 1 or 2...or I-1 or J+1...or N.

This kind of statistical method is also constructed on the basisoof two-valued logic. This logic is called two-valued logic of the interval.

5. A test study of the objective law of fuzzy subsets

In this section we shall introduce the statistical results of several fuzzy concepts.

In the sample surveys it is necessary to act according to the following rule: The surveyed people must have acquaintance with the concepts of fuzzy words and objective cases, in the meantime the surveyed people can show these concepts by using the various intervals.

Practice proves that there are statistical laws in the fuzzy concepts of language. That is to say, the statistical distribution of every sample set can be satisfied by the following characteristics.

(a) Under certain conditions the shapes of the statistical distributive curves which are obtained from different sample surveys of the same concept are about the same.

(b) Under certain conditions the areas of the statistical distributive curves which are obtained from different sample surveys of the same concept are about the same.

The handling of observations is in concordance with the following procedures:

(a) In the first place it is necessary to analyse the observations and give up some outcomes which do not seem to correspod with logic.

(b) After that a statistical distribution may be found by constructing the frequency table.

6. A new kind of frequency stability

The reason why there is a statistical law is that there is stability of membership frequency in the sample surveys.

In fuzzy set theory, a fuzzy subset $\underset{\sim}{B}$ may be substituted by an ordinary set denoted by B^*. Let A^* denote a real point within the fuzzy subset. The number of times that a real point A^* actually occurred throught N performances is called the frequency of real point A^* and is denoted P_i. The ratio P_i/N is called the relative frequency or membership frequency of the real point A^* in these N experiments and the ratio is denoted $F_n(A^*)$. Furthermore, a relative frequency is usually very unstable for small values of N, but it tends to stabilize as N increases. The number about which the relative frequency seems to stabilize is called the possibility of membership grade of the real point A^* with respect to the subset $\underset{\sim}{B}$. P_i/N represents the stability, and so does $\Sigma P_i/N$.

Example 6-1, Calculate the **frequency stability** of the total experimental sequence of the concept " young people ".

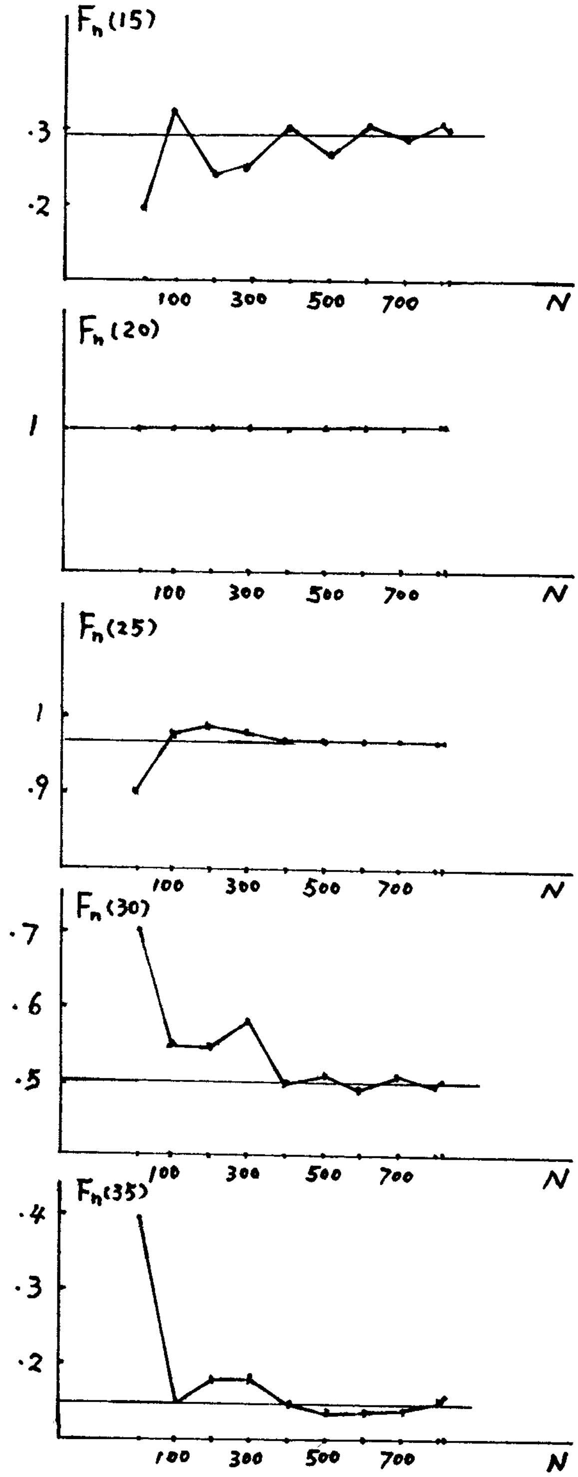
Seen as given in the following table 9-1.

A*		10	100	200	300	400	500	600	700	800	817
12	Pi	0	0	0	0	1	1	1	1	3	3
	Fn(A*)	0	0	0	0	.00	.00	.00	.00	.00	.00
13	Pi	0	1	1	2	3	3	3	3	6	6
	Fn(A*)	0	.01	.01	.01	.01	.01	.01	.00	.01	.01
14	Pi	0	9	11	15	27	31	55	57	70	70
	Fn(A*)	0	.09	.06	.05	.07	.06	.09	.08	.09	.09
15	Pi	2	32	50	78	118	133	186	205	246	248
	Fn(A*)	.2	.32	.25	.26	.30	.27	.31	.29	.31	.30
16	Pi	5	60	98	149	206	234	306	353	419	425
	Fn(A*)	.5	.6	.49	.50	.52	.47	.51	.50	.52	.52
17	Pi	6	68	118	176	238	272	346	402	468	475
	Fn(A*)	.60	.68	.59	.59	.60	.54	.58	.57	.59	.58
18	Pi	10	100	197	294	392	490	587	683	782	799
	Fn(A*)	1.0	1.0	.99	.98	.98	.98	.98	.98	.98	.98
19	Pi	10	100	198	295	394	492	590	687	786	803
	Fn(A*)	1.0	1.0	.99	.98	.99	.98	.98	.98	.98	.98
20	Pi	10	100	200	300	400	500	600	699	799	816
	Fn(A*)	1.0	1.0	1.0	1.0	1.0	1.0	1.0	1.0	1.0	1.0
21	Pi	10	100	300	399	499	598	697	797	797	814
	Fn(A*)	1.0	1.0	1.0	1.0	1.0	1.0	1.0	1.0	1.0	1.0
22	Pi	10	100	200	300	399	499	598	697	796	813
	Fn(A*)	1.0	1.0	1.0	1.0	1.0	1.0	1.0	1.0	1.0	1.0
23	Pi	10	100	200	299	398	498	596	695	794	811
	Fn(A*)	1.0	1.0	1.0	1.0	1.0	1.0	.99	.99	.99	.99
24	Pi	10	100	200	298	396	496	594	693	792	809
	Fn(A*)	1.0	1.0	1.0	.99	.99	.99	.99	.99	.99	.99
25	Pi	9	98	197	295	387	486	582	680	779	796
	Fn(A*)	.9	.98	.99	.98	.97	.97	.97	.97	.97	.97
26	Pi	9	76	154	234	290	373	437	522	603	620
	Fn(A*)	.9	.76	.77	.78	.73	.75	.73	.75	.75	.76

27	Pi	9	75	151	229	278	357	418	503	581	597
	Fn(A*)	.9	.75	.76	.76	.70	.71	.70	.72	.73	.73
28	Pi	9	75	149	225	267	345	404	489	565	581
	Fn(A*)	.9	.75	.75	.75	.76	.69	.67	.70	.71	.71
29	Pi	7	59	116	186	221	277	320	391	443	457
	Fn(A*)	.7	.59	.58	.62	.55	.55	.53	.56	.55	.56
30	Pi	7	55	110	174	201	253	291	356	403	415
	Fn(A*)	.7	.55	.55	.58	.50	.51	.49	.51	.50	.51
31	Pi	4	15	37	58	63	75	86	105	128	132
	Fn(A*)	.4	.15	.19	.19	.16	.15	.14	.15	.16	.16
32	Pi	4	15	37	58	63	75	86	105	128	132
	Fn(A*)	.4	.15	.19	.19	.16	.15	.14	.15	.16	.16
33	Pi	4	15	36	55	60	72	83	101	123	127
	Fn(A*)	.4	.15	.18	.18	.15	.14	.14	.14	.15	.16
34	Pi	4	15	36	55	60	72	82	100	122	126
	Fn(A*)	.4	.15	.18	.18	.15	.14	.14	.14	.15	.15
35	Pi	4	15	36	55	60	72	82	99	119	123
	Fn(A*)	.4	.15	.18	.18	.15	.14	.14	.14	.15	.15
36	Pi	1	2	3	5	6	7	10	11	15	15
	Fn(A*)	.1	.02	.02	.02	.02	.01	.02	.02	.02	.02
37	Pi	1	2	2	3	4	5	7	8	12	12
	Fn(A*)	.1	.02	.01	.01	.01	.01	.01	.01	.02	.01
38	Pi	1	2	2	2	3	4	6	7	11	11
	Fn(A*)	.1	.02	.01	.01	.01	.01	.01	.01	.01	.01
39	Pi	1	2	2	2	2	2	4	5	7	7
	Fn(A*)	.1	.02	.01	.01	.01	.01	.01	.01	.01	.01
40	Pi	0	0	0	0	0	0	2	3	5	5
	Fn(A*)	0	0	0	0	0	0	.00	.00	.01	.01
Σ Pi/N		15.7	13.9	13.7	13.8	13.4	13.2	13.2	13.3	13.5	13.5

Table 6-1: The frequency stability of the total experimental sequence.

For int·itive comprehension of the concept of frequency stability, the figures Fn(15), Fn(20), Fn(25), Fn(30), Fn(35), Pi/N are depicted in figure 9-1.



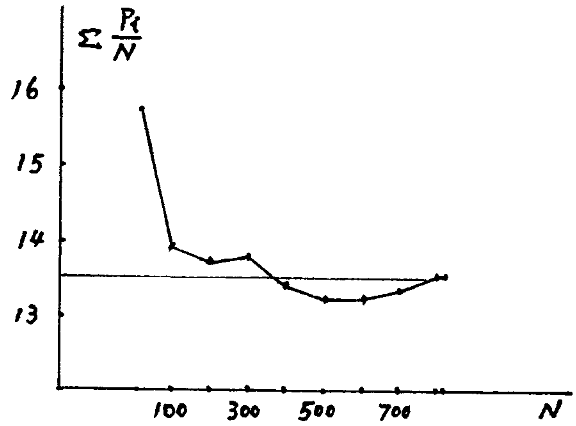


Figure 6-1: the frequency stability of Fn(15), Fn(20), Fn(25), Fn(30), Fn(35), Σ Pi/N.

7. Extensions of the concepts of random appearance

It is necessary to ask what is the difference between fuzzy and random appearances? Because fuzzy experiments satisfy all the characteristics of random experiments and there are statistical laws in the results of experiments, fuzzy appearances may thus be incorporated into the category of random appearances.

In finding membership functions the characteristics $\int_{-\infty}^{\infty} \mathcal{M}(x)dx > 1$ was discovered which must be the basis characteristic of the possibility distribution function.

Practical experience shows that the membership frequency shows some stability. However, the rigorous proof of this an important future task.

References: (Omission)