INSUFFICIENT DATA AND FUZZY AVERAGES

G.Sirbiladze, A.Sikharulidze

I.Vekua Institute of Applied Mathematics
Tbilisi State University
380043 University Street 2, Tbilisi, Georgia

(Received: 18.06.01; revised: 15.11.01)

Abstract

In this work three new versions of the Most Typical Value (MTV) of population are introduced - generalised weighted averages. The first version, WFEV_g, is a generalised version of the Weighted Fuzzy Expected Value (WFEV) for any fuzzy measure g on a finite set and, of course, it coincides with the WFEV used in sampling probability distribution. The second and third versions are the Weighted Fuzzy Expected Intervals WFEI and WFEI_g. These are generalisations of WFEV, the MTV-s of population respectively for sampling distribution and for any fuzzy measure g on a finite set, when the Fuzzy Expected Interval (FEI) exists but the Fuzzy Expected Value (FEV) does not. The work based on the Friedman-Schneider-Kandel (FSK) principle.

Key words and phrases: Fuzzy averages, weighted averages, fuzzy measure, Weighted fuzzy expected intervals.

AMS subject classification: 28E10.

1. *Introduction*

In the study of inexact data there are two classical approaches. When experimental data are exact enough, probabilistic-statistical methods are used to elaborate and estimate their general characteristics. If data presented are sufficiently inaccurate and have intervals, the methods of the theory of errors can be used successfully.

But, in some cases, when neither probabilistic-statistical methods nor those of the theory of errors give satisfactory results, one must, of course, search the nature of means of data reception (description, measurement, scaling, etc.), in order to find out the reason.

When data are represented in intervals and their distribution is obscure, they cover each other and are described or obtained by some person (insufficient expert data) participating in the process of obtaining or describing them, hence they become combined by nature. The so-called possibility uncertainty appears along with probabilistic-statistical uncertainty, which of course is produced by individual and calls for the application of fuzzy analysis methods.

In this case only probabilistic-possibility analysis will produce satisfactory results, which means using fuzzy methods to be explained below.

To obtain a general view of the set during the functional description of such data on the whole population, in many real situations it is impossible to observe the feature of additivity, which is unreliable and practically constitutes an additional limitation. In many cases it is more useful to use monotonic estimation instead of the additive kind to represent the human subjectivism (the study of subjective insufficient expert data).

For instance, consider three typical symptoms x_1, x_2, x_3, which indicate some disease y. Let the expert (doctor) provide objective-subjective data using his/her wide experience and the medical records of his/her patients (another expert would, of course, provide different data).

Assume, we have the following information: 80% of patients with disease y exhibit symptoms x_1 and x_2, 20% of them have symptoms x_1 and x_3. This information can be written using the additive, instead of the monotonic, measure g defined on the subsets (**Table 1.**),

$A \subseteq X$	g	g^*
$\{x_1\}$	0	1
$\{x_2\}$	0	0.8
$\{x_3\}$	0	0.2
$\{x_1, x_2\}$	0.8	1
$\{x_1, x_3\}$	0.2	1
$\{x_2, x_3\}$	0	1
$\{x_1, x_2, x_3\}$	0	1

Table 1. Distribution table showing dual measures g and g^*.

where g^* is called the dual measure of g. $g^*(A) = 1 - g(\overline{A})$. It must be said here that the dual measure contains the same information but codified in a different way.

Non-additive but monotonic measures were first used in fuzzy analysis in the 80s by M. Sugeno [11], and since the measure is connected with integral calculus, along with measurable functions, the measurable function integral was also constructed. This is called Sugeno's Integral for the compatibility function of the fuzzy subset with respect to the fuzzy measure - also known as FEV - which was then called fuzzy statistics by A.Kandel [5].

The fuzzy integral is the functional that relates some number, or compatibility value, to each measurable fuzzy subset when the fuzzy measure is already fixed. The fuzzy integral concept is presented along with that of the fuzzy measure: the possibility of condensing information when the fuzzy subset is estimated as the most typical compatibility value with respect to this measure. This is different from the probability mean even in the case when the probability measure is taken as a fuzzy measure because it is more "beneficial" than the average value.

In the present work the main estimators of fuzzy statistics of data processing are discussed, including the Fuzzy Expected Value (FEV) of population, the Fuzzy Expected Interval (FEI) and the Weighted Fuzzy Expected Value ($WFEV$) [4-6,8,11].

As already known, fuzzy means differ both in form and content from probability-statistical averages and other numeric characteristics, such as mode and median. Nevertheless, a coincidence does exist between "non-fuzzy" (objective) and "fuzzy" means in some cases [6]. For a given set of fuzzy subset compatibility function values from interval [0;1], the fuzzy mean distinguishes the most typical characteristic compatibility value (FEV) or interval of compatibility values (FEI).

Fuzzy statistics play an essential part in probability-possibility analysis and they are used very effectively in fuzzy expert (decision-making) systems. In the case of fuzzy data, fuzzy means are mainly built on population groups (FEV) and if these data are insufficient the FEI will replace the FEV.

It is important to mention that the FEV seldom satisfies demands on the most typical value (MTV). In the case of sampling distribution of population M. Friedman, M. Schneider and A. Kandel constructed a process for calculating the Weighted Fuzzy Expected Value ($WFEV$), which is based on a principle with two postulates (FSK). This value is viewed by these well-known specialists as the most typical value of population. ($MTV = WFEV$).

Software was created for estimating Weighted Values, as well as for calculating the FEV and FEI.

Practically speaking, $WFEV_g$ is a calculation process using probabilistic representations on a finite set, the so-called class of associated probability distributions [1], which enables one to estimate associated probabilities - fuzzy measure values -by intervals of belief when the representation of data is inexact. This means that it is possible to represent (estimate) the fuzzy measure by intervals, which is the usual attribute of interval extension in $WFEI$ or $WFEI_g$. Thus, in this case one does not face the problem of uncertain fuzzy distribution. The authors believe that the use of the $WFEI$ is a perspective that needs further research, which would create new perspectives of fuzzy data processing when data are insufficient and obscure.

2. *Fuzzy Measure and the FEV*

Definition 1 [6]: Let (X, F) be a measurable space, let F be a Borel field (σ-algebra), $g : F \Longrightarrow [0,1]$ function is called a fuzzy measure if the following is true:

(i) $g(\oslash) = 0$, $g(X) = 1$;

(ii) If $A \subset B$ then $g(A) \leq g(B)$;

(iii) If $\{A_k / 1 < k < \infty\}$ is a monotonic sequence, $\forall A_k \in F$, then $\lim\limits_{k \to \infty} g(A_k) = g\left(\lim\limits_{k \to \infty} A_k\right)$.

(X, F, g) is called a fuzzy measure space.

Let $\chi_{\widetilde{A}}$ be the compatibility function of the fuzzy subset $\widetilde{A}$ and $\chi_{\widetilde{A}} : X \to [0,1]$ is a F-measurable function, i.e. $\forall T \in [0,1] : H_T = \{x \in X / \chi_{\widetilde{A}}(x) \geq T\} \in F$.

Definition 2 [6]: FEV of a compatibility function $\chi_{\widetilde{A}}$ of the fuzzy subset $\widetilde{A}$ with respect to the fuzzy measure g is Sugeno's integral over:

$$FEV(\chi_{\widetilde{A}}) = \int_X \chi_{\widetilde{A}} \circ g(\cdot) \equiv \sup_{T \in [0,1]} \{T \wedge g(H_T)\} \qquad (2.1)$$

where $\wedge$ indicates a minimum of two arguments.

If $g(H_T)$, $T \in [0,1]$ is a continuous function, then the geometric interpretation of the FEV is as shown below (Figure 1):

Clearly, the FEV somehow "averages" values of the compatibility function $\chi_{\widetilde{A}}$, although not with respect to the statistical mean but by cutting the subsets of level T values of fuzzy measure g of which "fuzzy weights" are sufficiently "high" or sufficiently "low".

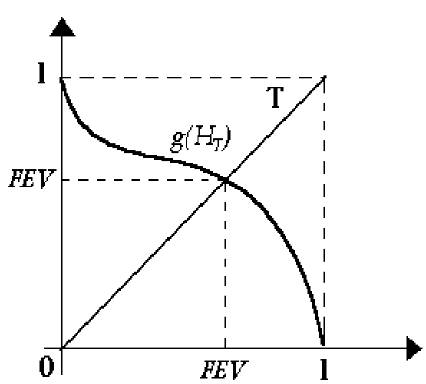

Figure 1. Geometric concept of calculating the FEV

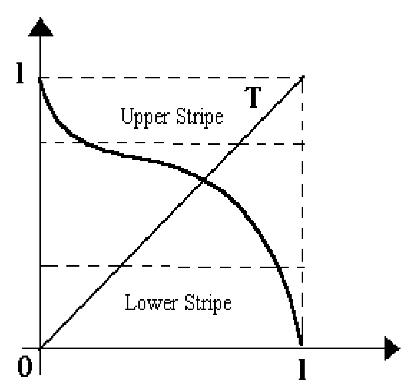

Figure 2. FEV cuts "upper" and "lower" strips of $g(H_T)$

Thus, the FEV gives that concrete possible value of compatibility function $\chi_{\widetilde{A}}$, this being the most typically characteristic among all possible values with respect to the fuzzy measure g, which is obtained by cutting the "upper" and "lower" stripes on the graph of $g(H_T)$ (Figure 2).

This is a condensation of information given in the FEV by $\chi_{\widetilde{A}}$ and g , which is the Most Typical Value (MTV) of all compatibility values.

Consider the situation where $X = \{x_1, x_2, ..., x_n\}$ is a finite set arranged in the following way: $\chi_{\widetilde{A}}(x_1) \leq \chi_{\widetilde{A}}(x_2) \leq \cdots \leq \chi_{\widetilde{A}}(x_n)$. Denote: $X_i = \{x_i, ..., x_n\}, i = 1, 2, .., n$. As known, the FEV can be calculated so [10]:

$$FEV = \max_i \left\{\chi_{\widetilde{A}}(x_i) \wedge g(X_i)\right\} = \min_i \left\{\chi_{\widetilde{A}}(x_i) \vee g(X_i)\right\}, \qquad (2.2)$$

where $\vee$ - is a maximum of two arguments. If $\chi_i \equiv \chi_{\widetilde{A}}(x_i)$, $g_i \equiv g(X_i)$, then the possible geometric interpretation of equation (2) is as shown below (Figure 3):

Some interesting examples concerning the calculation of the FEV will be considered below.

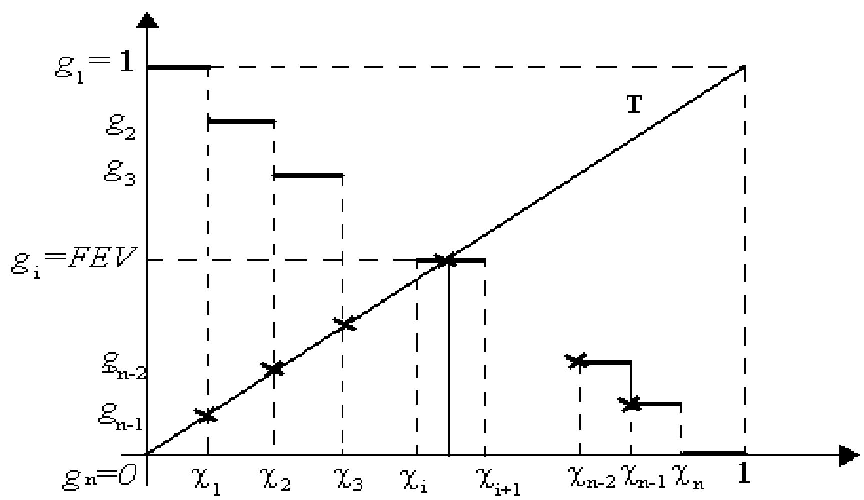

Figure 3. FEV - Discrete case. x - indicates $\chi_i \wedge g_i$ value, maximum of which is FEV.

Example 1. [5] The following statistical data were gathered in Bier-Sheva, Israel. During 55 years since 1920 the maximum temperatures registered there on July 1st were the following:

51 days	90°F-92F (average 91°F)
1 day	106°F,
1 day	122°F,
1 day	124°F,
1 day	132°F,

The problem is to determine what is the temperature of hot weather in this city on July 1st? And what temperature characterises hot weather in Bier-Sheva on this particular day?

The base variable "hot weather" is of course the fuzzy subset of the temperature distribution on the whole population. For one assessor living in the South the temperature of hot weather is more than 80°F, for another assessor living in North hot weather is a temperature somewhere below 80°F. This is the reason why the notion "hot weather" is fuzzy and is given by the function constructed by some expert. Suppose the compatibility curve is as shown below (Figure 4):

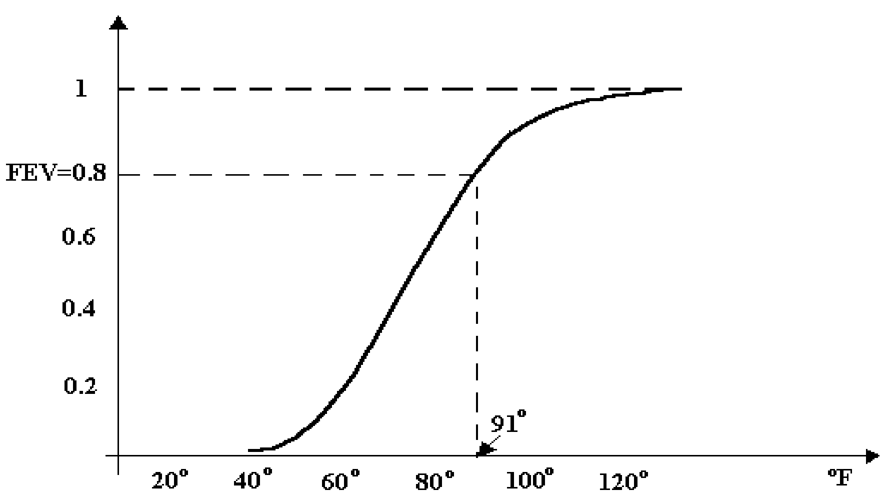

Figure 4. Compatibility curve for "hot" weather".

To solve this problem classic statistics will be used at first. Probability $mean = (91{\cdot}51{+}484)/55 = 93,2°F$, $median = 91°F$. Clearly, *mean* cannot describe the typically characteristic temperature of "hot weather" on July 1st because it must coincide with the *median* (high temperatures vary from 90°F to 92°F with a higher frequency). What does Fuzzy Statistics have to offer? The FEV. If uniform distribution is used in the case of fuzzy measure g (there is not any other information available about g) then $g(H_T) = card(H_T)/55$, where $card$ is the cardinality of set H_T. The FEV is calculated using equation (2): $FEV(\chi_{\tilde{A}}) = 0.8$ which means temperature $\chi_{\tilde{A}}^{-1}(0.8) = 91°F$, i.e. according to the expert, who assesses "hot weather" by the compatibility curve shown in Fig.5, the most typically characteristic temperature of hot weather on July 1st is 91°F.

If the expert is changed and his/her compatibility function is more "southern" (Figure 5), then $FEV(\chi_{\tilde{A}}) = 0.01543, = 110°F, \chi_{\tilde{A}}^{-1}(0.01543) = 110°F$ wherever $mean(\chi_{\tilde{A}}) = 0.0235$, i.e. $mean = 94°F$, which in reality is very "low" according to the southern expert.

It can be said that the FEV is a subjective, expert characteristic for population; the most typically characteristic value among the compatibility values of the fuzzy subset according to the aforementioned expert.

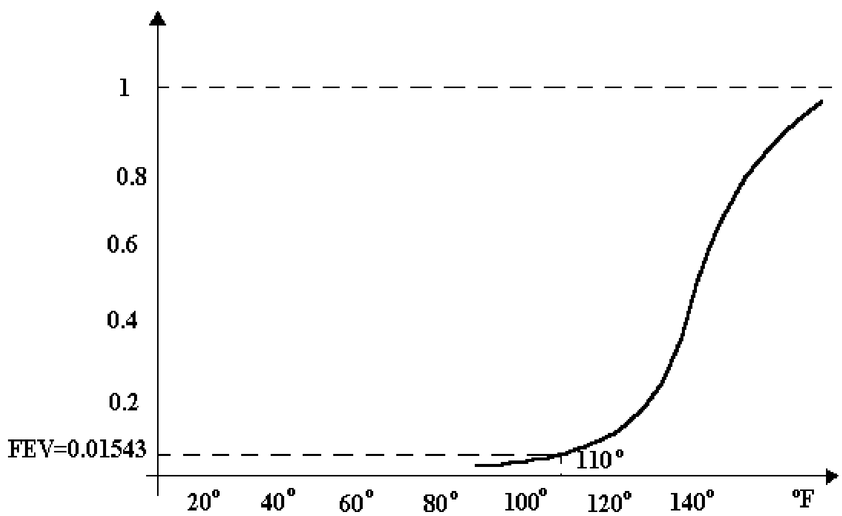

Figure 5. Compatibility curve for "hot weather" ("southern")

Example 2. Let the base variable be "high salary", which creates some fuzzy subset on the set of employees. Consider the salary earned by a number of people and the subjective (expert) compatibility values for χ shown in the following table:

1 person earns 3.00 $\rightarrow \chi = 0.40$
3 person earns 4.00 $\rightarrow \chi = 0.50$
4 person earns 4.20 $\rightarrow \chi = 0.55$
2 person earns 4.50 $\rightarrow \chi = 0.60$
2 person earns 10.00 $\rightarrow \chi = 1.00$

Suppose that the following statistical data are available to calculate the FEV:

# of group	x_i	n_i	χ_i	$n^{(i)}$	$g_i = n^{(i)}/n$	$\chi_i \wedge g_i$
1	3.00	1	0.4	12	1	0.4
2	4.00	3	0.5	11	11/12	0.5
3	4.20	4	0.55	8	8/12	0.55
4	4.50	2	0.6	4	4/12	0.33
5	10.00	2	1.0	2	2/12	0.16

where n_i - is the number of people in i-th group; $n^{(i)} \equiv \sum_{j=1}^{n} n_j$, $i = 1, 2, ..., n; n = 5.$;

As in the previous example, the sampling distribution for the fuzzy measure g on whole population is taken. Then $FEV = 0.55$, which coincides with the median (Kandel showed that for unimodal variational sampling if the fuzzy measure has sampling distribution, the FEV coincides with the median) $\chi_{\tilde{A}}(0.55) = 4.2$; i.e. a typical high salary on the whole population $= 4.2$.

When receiving data in extreme situations the FEV does not provide a "logical" expected value because it is assumed that in this case the information available on the population is insufficient. Consider the following:

Example 3. Let the compatibility function for the variable "old" be

$$\chi(x) = \begin{cases} 0, & x < 0 \\ x/100, & 0 \le x \le 100 \\ 1, & x > 100 \end{cases},$$

and the statistical distribution of population groups be

10 people are [10-20] years old,
25 people are 30 years old,
15 people are 40 years old,
35 people are [45-55] years old,
20 people are [60-70] years old.

As in Example 2, the table of statistical data is as shown below:

# of group	x_i	n_i	χ_i	$n^{(i)}$	$g_i = n^{(i)}/n$
1	[10;20]	10	[0.1;0.2]	100	1.00
2	30	25	0.3	90	0.90
3	40	15	0.4	55	0.65
4	[45;55]	35	[0.45;0.55]	50	0.50
5	[60;70]	20	[0.6;0.7]	20	0.20

It is clear that the FEV cannot be calculated with this data, and if the same is done as in Example 1 (when for interval [90°F-92°F] an average of 91°F was taken) the result will be unsatisfactory because information will be lost and this will lead to a significant reduction in informational entropy.

By introducing interval algebra M. Schneider and A. Kandel [8] offer a new way in which operations of $\wedge$ (minimum) and $\vee$ (maximum) are defined on intervals, and the procedure for calculating the FEV (on a finite set) is generalised. This method is called the Fuzzy Expected Interval (FEI).

3. *Fuzzy Expected Interval (FEI)*

The concept of the FEI as a method was developed to overcome inaccurate fuzzy information when calculating the FEV. Naturally, the FEI must give the same results as the FEV when the intervals are one-point sets and display stability concerning the FEV in the case of intervals with a "small" length, which is used to define $\vee$ and $\wedge$ operations in interval algebra. Let us use the definitions and results from [8] (without proof):

Definition 3. If $S = [\underline{s}, \overline{s}]$and $R = [\underline{r}, \overline{r}]$ are intervals, then

$$\begin{array}{l} \max\{S,R\} = S \quad if \quad \forall s \in S : \exists \widetilde{r} \in R \quad such \quad that \quad s > \widetilde{r}, \\ \min\{S,R\} = S \quad if \quad \forall s \in S : \exists \widetilde{r} \in R \quad such \quad that \quad s < \widetilde{r}. \end{array} \tag{3.3}$$

Proposition 1. If $S \cap R = \emptyset$ then

$$\begin{array}{c} \max\{S,R\} = \left\{ \begin{array}{lll} R & if & \underline{r} > \overline{s} \\ S & if & \underline{s} > \overline{r} \end{array} \right., \\ \min\{S,R\} = \left\{ \begin{array}{lll} R & if & \overline{r} < \underline{s} \\ S & if & \overline{s} < \underline{r} \end{array} \right.. \end{array} \tag{3.4}$$

Proposition 2. If $S \cap R = \emptyset$, $S \subseteq R$ and $R \nsubseteq S$, then

$$\begin{array}{c} \max\{S,R\} = \left\{ \begin{array}{lll} R & if & \overline{r} > \overline{s} \\ S & if & \overline{s} > \overline{r} \end{array} \right., \\ \min\{S,R\} = \left\{ \begin{array}{lll} R & if & \overline{s} > \overline{r} \\ S & if & \overline{r} > \overline{s} \end{array} \right.. \end{array} \tag{3.5}$$

Definition 4. Suppose $S \subseteq R$, then $\exists T (T = [\underline{t}; \overline{t}])$ so that

$\max\{S,R,T\} = T$ if $\forall t \in T : \exists \widetilde{s} \in S$, such that $t \geq \widetilde{s}$ and $\exists \widetilde{r} \in R$, such that $t \geq \widetilde{r}$;

$\min\{S,R,T\} = T$ if $\forall t \in T : \exists \widetilde{s} \in S$, such that $t \geq \widetilde{s}$ and $\exists \widetilde{r} \in R$, such that $t \leq \widetilde{r}$.

Proposition 3. If $R \subseteq S$, then

$$\max\{R,S\} = [\underline{r};\overline{s}]\,, \min\{R,S\} = [\underline{s};\overline{r}]\,. \tag{3.6}$$

Definition 5. Suppose R and S are any intervals from $\Im(\Re)$ (sets of all intervals on real numbers $\Re$). One can say that S is "higher" then R if $\overline{s} \geq \overline{r}$.

Thus, there is a possibility to define $\wedge$ and $\vee$ operations on any interval. Now, example 3 can be concluded as shown below:

$$FEI = \max\{[0.1;0.2], 0.3, 0.4, [0.45;0.5], 0.2\} = [0.45;0.5],$$

where $[0.1;0.2] = min\{[0.1;0.2],1\}, 0.3 = min\{0.3, 0.9\}, 0.4 = min\{0.4, 0.65\}$, $[0.45;0.5] = min\{[0.45;0.55], 0.5\}$, $0.2 = min\{[0.6;0.7], 0.2\}$, but $\chi^{-1}([0.45;0.5]) = [45;50]$. That is to say, the most typical age in a given population regarding the variable "old" is the interval [45-50].

Since there are some examples in which the information available for the frequency distribution of the population is scarce and inaccurate, the frequencies of groups are given by intervals.

Example 4. Consider the base variable "old" with the same compatibility function as in example 3. The population consists of two groups:

# of group	x_i	n_i	χ_i	$n^{(i)}$	g_i
1	15	[10;15]	0.15	?	?
2	20	[20;30]	0.20	?	?

For instance, this means that in the first group 10 to 15 children are fifteen years old, and in the second group 20 to 30 children are twenty years old. What is the MTV?

Generally speaking, the values of fuzzy measure g_i are intervals whose upper and lower boundaries are calculated as follows [8]:

$$\begin{aligned} \underline{g}_j &= \frac{\sum\limits_{i=1}^{k} \min\{\underline{n}_i;\overline{n}_i\}}{\sum\limits_{i=j}^{k} \min\{\underline{n}_i;\overline{n}_i\} - \sum\limits_{i=1}^{j-1} \max\{\underline{n}_i;\overline{n}_i\}}, \\ \overline{g}_j &= \frac{\sum\limits_{i=1}^{k} \max\{\underline{n}_i;\overline{n}_i\}}{\sum\limits_{i=j}^{k} \max\{\underline{n}_i;\overline{n}_i\} - \sum\limits_{i=1}^{j-1} \min\{\underline{n}_i;\overline{n}_i\}}, \end{aligned} \tag{3.7}$$

where k is the number of groups in the whole population and $[\underline{n}_i;\overline{n}_i] \equiv n_i$ are frequency intervals of i-th group. If formulae (7) from example 4 are used, $g_i = \left[\underline{g}_i;\overline{g}_i\right]$ intervals, where $i = 1,2$, will be calculated so that $\underline{g}_1 = \overline{g}_1 = 1$, $\underline{g}_2 = 20/(10+30) = 0.25$, $\overline{g}_2 = 30/(10+30) = 0.75$. And the following table will be obtained:

# of group	x_i	n_i	χ_i	$n^{(i)}$	g_i
1	15	[10;15]	0.15	[30;45]	[1;1]
2	20	[20;30]	0.20	[20;30	[0.25;0.75].

Then $FEI = \max\{\min(0.15, 1), \min(0.2, [0.25; 0.75])\} = \max\{0.15, 0.2\} = 0.2$, but $\chi^{-1}(0.2) = 20$, which means that the most typical group in the whole population is the second one.

In many cases the information is more uncertain than in the aforementioned examples and is represented by the so-called "linguistic variables". These are "about", "more or less", "more", "much more", etc. In every problem the subject (expert) constructs a table of relationships for each indicative variable of the population ("person" in the present case), which transfers linguistic variables to the frequency intervals (mapping table):

Linguistic Variable	Lower Border	Upper Border
Almost	$x - 10\%$	$x - 1\%$
More or less	$x - 10\%$	$x + 10\%$
Much more	$2x$	$+\infty$

Notice that while receiving data, each linguistic variable creates some population group with a frequency interval. In this case the FEI has already been calculated.

One example of how to calculate the FEI by means of one expert system of decision-making is given below. In this example the general system of decision-making is as follows:

"If the condition is fulfilled, then act".

Consider the situation for the population when a decision must be made regarding a raise of salary.

If "high income", then "raisesalary".

More concretely:

If the salary earned is ≥ 5, then it must be raised by 1%.

Suppose the information on population groups is as follows:

More or less 30 people earn \$2.5,
50 people earn \$[4-5],
70-100 people earn \$5.5,
50-70 people earn \$[7-8].

The question arises: Does this population of employees receive a raise?

Let the compatibility function of the base variable "high salary" be as follows:

$$\chi(x) = \begin{cases} 0, & x < 0, \\ x/10, & 0 \le x \le 10, \\ 1, & x > 10. \end{cases}$$

The first population group is created by the linguistic variable "more or less". The following distribution interval will be obtained from the above-mentioned mapping table:

$$[30 - 10\% \text{ of } 30; 30 + 10\% \text{ of } 30] = [27, 33],$$

and the following distribution table will be obtained:

# of group	$[\underline{x}_i;\overline{x}_i]$	$n_i = [\underline{n}_i;\overline{n}_i]$	$\chi_i = \left[\underline{\chi}_i;\overline{\chi}_i\right]$	$g_i = \left[\underline{g}_i;\overline{g}_i\right]$
1	2.5	27-33	0.25	1
2	4.0-5.0	50	0.4-0.5	0.84-0.89
3	5.5	70-100	0.55	0.55-0.68
4	7.0-8.0	50-70	0.7-0.8	0.24-0.28

Then $FEI = \max\{\min(0.25, 1), \min([0.4; 0.5], [0.84; 0.89]), \min(0.55, [0.55; 0.68]), \min([6.7; 0.8], [0.24; 0.28])\} = \max\{0.25, [0.4; 0.5], 0.55, [0.24; 0.9]\}$ or $FEI = 0.55$, but $\chi^{-1}(FEI) = \chi^{-1}(0.55) = 5.5 = \chi^{-1}(MTV)$. Because the $\chi^{-1}(MTV) > 5.05$, one can say that employees get a raise.

Despite the fact that the FEV gives a good representation of the Most Typical Population Group ($MTPG$) (when there are sufficient data) and the FEI gives an interval estimation of the MTV of the compatibility curve (when there are insufficient data on population groups), there are some cases when both of them give unsatisfactory results. Consider the examples shown below:

Example 5. Suppose that the following table of compatibility values is obtained:

# of group	n_i	χ_i	g_i	$\max(\chi_i, g_i)$
1	70	0.05	1	0.05
2	30	0.3	0.3	0.3

If one chooses $FEV = 0.3$as the most characteristic value of function χ, then the group of 70% with compatibility value 0.05 is ignored. The *mean* $= 0.125$ is also unsatisfactory. It would be better to acknowledge two facts when calculating the MTV [4]:

1. The MTV must consider groups with a higher frequency in the whole population;
2. The MTV must consider closeness with groups with high compatibility values.

Note that these factors are conditional and vary according to the subjective opinion about the MTV. But it should be said in advance that these two factors play an essential role in the elaboration of a new method around the FEI.

4. *Weighted Fuzzy Expected Value* ($WFEV$)

M.Friedman, M.Schneider and A.Kandel offered a new scheme for calculating the MTV [4], which is based on a two-factor principle: Taking, for example, the following two population groups:

# of group	χ	n
i	χ_i	n_i
j	χ_j	n_j

Suppose $n_i > n_j$, then:

1. 1. Population effectiveness: the MTV will be 'less far' from χ_i than from χ_j since $n_i > n_j$.

   2. The effective location of the MTV with respect to compatibility values: The distance between the MTV and the compatibility value of i-th group $|\chi_i - MTV|$ will participate in the definition of the MTV with a "low" weight, as "large" this distance might be. This weight will be proportional to $w\left(|\chi_i - MTV|\right)$, where w is some strictly decreasing function.

Suppose a variational sampling $\sim \begin{pmatrix} (x_1, x_2, ..., x_k) \\ (n_1, n_2, ..., n_k) \end{pmatrix}$ is given, $\chi_i = \chi_{\widetilde{A}}(x_i)$ are compatibility values of some fuzzy set $\widetilde{A} \subset X = \{x_1, x_2, ..., x_k\}$, $w(x)$ is a non-negative monotonically decreasing function defined over the interval [0,1] and $l > 1$ is a real number. Consider the following equation with respect to s:

$$s = \frac{\chi_1 w\left(|\chi_1 - s|\right) n_1^l + \chi_2 w\left(|\chi_2 - s|\right) n_2^l + \cdots + \chi_k w\left(|\chi_k - s|\right) n_k^l}{w\left(|\chi_1 - s|\right) n_1^l + w\left(|\chi_2 - s|\right) n_2^l + \cdots + w\left(|\chi_k - s|\right) n_k^l}. \quad (4.8)$$

Definition 6. The solution of equation (8) is called the Weighted Fuzzy Expected Value ($WFEV$) of order l with the attached weight function w of compatibility values $(\chi_1, ..., \chi_k)$. $\left(MTV \equiv WFEV\left(\chi_{\tilde{A}}, w\right)\right)$.

The Parameter l measures the dependence of frequencies of population groups on the $WFEV$. The speed at which function w decreases defines the "closeness" of $WFEV$ to higher compatibility values of χ_i. With the above-mentioned principle, which consists of two factors, the mapping of the weighting invariant to the $WFEV$ follows from definition 8, with the MTV being a fixed mapping point. The authors of [11] use the function $w(x) = e^{-\lambda x}(\lambda > 0)$ instead of w. Specifically, for a pair (l,λ) values $l = 2, \lambda = 1$. To solve equation (8) they use the iteration method $s_n = f(s_{n-1})$, where $s_0 = FEV$ (the function f is the value in the right hand side of equation (8)), and after 3-4 steps they achieve an accuracy of $\varepsilon = 10^{-3}$. A discussion on some of the examples of the use of the $WFEV$ follows: In the case of example 5, if $l = 2$ and $\lambda = 1$, then $WFEV \approx 0.083$, $FEV = 0.3$, $mean = 0.125$, $median = 0.05$. Clearly, the FEV and $median$ ignore groups with 70% and 30% density accordingly. The $mean$ is close to the compatibility value with a higher density but represents a more insufficient measure of typicality than the $WFEV$. The latter uses a two-factor principle and is the most typical value for the population. According to the authors of [4] $(MTV = WFEV)$.

Example 6. The population consists of two groups with the following table of compatibility values:

# of group	χ_i	n_i	g_i
1	0.125	7	1
2	0.375	19	0.93
3	0.625	31	0.74
4	0.875	43	0.43

If $l = 2$ and $\lambda = 1$, then $FEV = 0.625$, $mean = 0.65$, $median = 0.625$, $mode = 0.875$, $WFEV = 0.745$. As in the previous example, the $mean$ is a "better" MTV than $FEV = median$, but "worse" than $mode = 0.875$. This is best summarised as $WFEV$, so $MTV = WFEV$.

Example 7. The population consists of three groups with the following table of compatibility values:

# of group	χ_i	n_i	g_i
1	0.2	35	1
2	0.3	25	0.65
3	0.6	40	0.4

Then $FEV = 0.4$, $mean = 0.385$, $median = 0.3$, $WFEV = 0.402$, $(l = 2, \lambda = 1)$, $mode = 0.6$.

Clearly, neither mean nor median are sufficient MTVs. The mean is a little bit better than median, the FEV is better than *mean* and the $WFEV$ is much better than both, because it is moved close to the compatibility value of higher density group and also considers the existence of first and second groups with 60% density.

5. *Weighted Fuzzy Expected Interval ($WFEI$)*

It is important to note that it is impossible to calculate the FEV when data on population groups are insufficient. Hence, a method for calculating the FEI was elaborated, which effectively uses two operations: $\vee$ and $\wedge$ from interval algebra. This procedure is stable and for one-point intervals the FEI coincides with the FEV. Naturally, the same problem arises during the calculation process of the $WFEV$ when the starting point of iteration process $s_n = f(s_{n-1})$ cannot be found. But the FEI does exist. How can the FEI be used to build similar process? An attempt to construct a new iteration process using interval analysis will be made, where the essential base component will be the FEI using principles for constructing the $WFEV$.

Suppose a variational sampling $\sim \begin{pmatrix} (x_1, x_2, ..., x_k) \\ (n_1, n_2, ..., n_k) \end{pmatrix}$ is given, $\chi_i = \chi_{\widetilde{A}}(x_i)$ are compatibility values of some fuzzy set $\widetilde{A} \subset X = \{x_1, x_2, ..., x_k\}$. n_i and χ_i are intervals: $n_i = [\underline{n}_i; \overline{n}_i], \chi_i = \left[\underline{\chi}_i; \overline{\chi}_i\right]$, $i = 1, 2, ..., k$. Let $w(x)$ be a non-negative monotonically decreasing function defined over the interval $[0,1]$and $l > 1$ be a real number.

Definition 7. The Weighted Fuzzy Expected Interval ($WFEI$) of order l with attached weight function w of compatibility values $\{\chi_1, ..., \chi_k\}$ is called the limit of the iteration process of the combinatorial interval extension:

$$s_n = \frac{\sum\limits_{i=1}^{k} w\left(\left|\left[\underline{\chi}_i; \overline{\chi}_i\right] - s_{n-1}\right|\right) \cdot \left[\underline{n}_i^l; \overline{n}_i^l\right] \cdot \left[\underline{\chi}_i; \overline{\chi}_i\right]}{\sum\limits_{i=1}^{k} w\left(\left|\left[\underline{\chi}_i; \overline{\chi}_i\right] - s_{n-1}\right|\right) \cdot \left[\underline{n}_i^l; \overline{n}_i^l\right]} \tag{5.9}$$

where $s_0 \equiv FEI$.

It is denoted by $WFEI\left(\chi_{\widetilde{A}}, w\right)$. It is clear that $WFEI$ is the interval extension of $WFEV$, when FEV does not exist, but FEI [4] exists.

An essential proposition, which unifies all known weighted means presented in this paper and retains correctness of generalization of statistical notions, will be stated below:

Proposition 4. (without proof) If $FEV = FEI$, intervals of compatibility values χ_i and frequencies n_i are one-point intervals then

$$WFEV\left(\chi_{\widetilde{A}}, w\right) = WFEI\left(\chi_{\widetilde{A}}, w\right).$$

Note that for the convergence of iteration process equation (9) the property of compression of the function w is sufficient.

6. *Weighted Fuzzy Expected Value with Respect to Fuzzy Measure ($WFEV_{gl}$)*

Suppose a variational sampling $\sim \begin{pmatrix} (x_1, x_2, ..., x_k) \\ (n_1, n_2, ..., n_k) \end{pmatrix}$ is given, $\chi_i = \chi_{\widetilde{A}}(x_i)$ are compatibility values of some fuzzy set $\widetilde{A}$. Let w be a non-negative monotonically decreasing function defined over the interval [0,1] and $l > 1$ be a real number.

The equation (8) can be written in the following way:

$$s = \frac{\chi_1 w\left(|x_1 - s|\right)\left(\frac{n_1}{k}\right)^l + \chi_2 w\left(|x_2 - s|\right)\left(\frac{n_2}{k}\right)^l + \cdots + \chi_k w\left(|x_k - s|\right)\left(\frac{n_k}{k}\right)^l}{w\left(|x_1 - s|\right)\left(\frac{n_1}{k}\right)^l + w\left(|x_2 - s|\right)\left(\frac{n_2}{k}\right)^l + \cdots + w\left(|x_k - s|\right)\left(\frac{n_k}{k}\right)^l}. \quad (6.10)$$

Definition 8. Fuzzy measure, which, for any subset A of sampling $X = \{x_1, ..x_k\}$, is equal to l-th power of frequency of A

$$g^l_{sampling}(A) \equiv \left(\frac{\sum\limits_{x_i \in A} n_i}{k}\right)^l = \left(\frac{n_A}{k}\right)^l,$$

is called a fuzzy measure induced with sampling distribution of l-th power.

Then

$$g^l_{sampling}(\{x_i\}) = \left(\frac{n_i}{k}\right)^l, i = 1, 2, ..., k. \quad (6.11)$$

Obviously, during weighting, values of measure $g^l_{sampling}$ in equation (10) only on sets of one element (fuzzy "weights" of sets of one element) are considered.

Let $X = \{x_1, ..., x_k\}$ be a finite set, let $\left(X, 2^X, g\right)$ be a fuzzy measure space, let $\chi_{\widetilde{A}}$ be a compatibility function of fuzzy subset $\widetilde{A}$, $\chi_{\widetilde{A}} : X \to [0; 1]$ $\left(\chi_i = \chi_{\widetilde{A}}(x_i)\right)$; let w be some "weight" function and let $l > 1$ be a real number.

Considering the equation (10) and definition 8 it is possible to compose two new postulates of constructing $MTV = WFEV$ with respect to the fuzzy measure g on the set X, which will be later called Friedman-Schneider-Kandel principles (FSK):

1. Fuzzy measure distribution effectiveness: MTV will be "less far" from χ_i than from χ_j if $g(\{x_i\}) > g(\{x_j\})$.

2. The effective location of MTV with respect to compatibility values: The distance between MTV and compatibility values χ_i (of the element $x_i \in X$):$|\chi_i - MTV|$ will participate in the definition of the MTV with a "low" weight, as "large" this distance might be. This weight will be proportional to $w(|\chi_i - MTV|)$, where w is some strictly decreasing function.

Similarly to equation (10) and equation (11) consider the following equation with respect to s:

$$s = \frac{\sum\limits_{i=1}^{k} \chi_i w\left|(\chi_i - s)\right| g^l(\{x_i\})}{\sum\limits_{i=1}^{k} w\left|(\chi_i - s)\right| g^l(\{x_i\})}. \tag{6.12}$$

Definition 9. The solution of equation (12) is called the Weighted Fuzzy Expected Value of order l with the attached weighted function w of compatibility function χ with respect to the fuzzy measure g.

It is denoted by $WFEV_{gl}\left(\chi_{\widetilde{A}}, w\right)$, $(MTV = WFEV_{gl})$.

On the set $X = \{x_1, ..., x_k\}$ there exist $k!$ permutations. Denote any permutation by $\sigma = (\sigma(1), \sigma(2), ..., \sigma(k))$, the set of all possible permutations by S_k

Definition 10 [1]. If $\sigma \in S_k$ is any permutation, then the following probability distribution

$$P_\sigma^{(l)}\left(x_{\sigma(1)}\right) = g^l\left(\left\{x_{\sigma(1)}\right\}\right),$$
$$P_\sigma^{(l)}\left(x_{\sigma(2)}\right) = g^l\left(\left\{x_{\sigma(1)}, x_{\sigma(2)}\right\}\right) - g^l\left(\left\{x_{\sigma(1)}\right\}\right),$$
$$\cdots\cdots\cdots\cdots$$
$$P_\sigma^{(l)}\left(x_{\sigma(i)}\right) = g^l\left(\left\{x_{\sigma(1)}, ..., x_{\sigma(i)}\right\}\right) - g^l\left(\left\{x_{\sigma(1)}, ..., x_{\sigma(i-1)}\right\}\right),$$
$$\cdots\cdots\cdots\cdots$$
$$P_\sigma^{(l)}\left(x_{\sigma(k)}\right) = 1 - g^l\left(\left\{x_{\sigma(1)}, ..., x_{\sigma(k-1)}\right\}\right).$$

is called an associated probability distribution of the fuzzy measure g^l; $\left\{P_\sigma^{(l)}\right\}_{\sigma \in S_k} = \left\{P_\sigma^{(l)}\left(x_{\sigma(1)}\right), ..., P_\sigma^{(l)}\left(x_{\sigma(k)}\right)\right\}_{\sigma \in S_k}$ is called the class of associated probabilities of the fuzzy measure g^l.

It is known that for $\forall x_i \subset X$ set, $\exists \tau_i \in S_k$ permutation such that

$$g(\{x_i\}) = P_{\tau_i}^{(l)}(x_i) \equiv P_{\tau_i}^{(l)}\left(x_{\tau_i(1)}\right).$$

Then equation (12) will take the following form:

$$s = \frac{\sum\limits_{i=1}^{k} \chi_i w \left|(\chi_i - s)\right| P_{\tau_i}^{(l)}\left(x_{\tau_i(1)}\right)}{\sum\limits_{i=1}^{k} w \left|(\chi_i - s)\right| P_{\tau_i}^{(l)}\left(x_{\tau_i(1)}\right)}. \qquad (6.13)$$

This is the probability representation of $WFEV_{gl}$ by associated probabilities $P_{\tau_1}^{(l)}, P_{\tau_2}^{(l)}, ..., P_{\tau_n}^{(l)}$ of the fuzzy measure g.

Obviously, we can construct the iteration process for equation (13) as we did for definition 8:

$$s_n = \frac{\sum\limits_{i=1}^{k} \chi_i w \left|(\chi_i - s_{n-1})\right| P_{\tau_i}^{(l)}\left(x_{\tau_i(1)}\right)}{\sum\limits_{i=1}^{k} w \left|(\chi_i - s_{n-1})\right| P_{\tau_i}^{(l)}\left(x_{\tau_i(1)}\right)}$$

where $s_0 = FEV\left(\chi_{\widetilde{A}}\right)$.

Let χ_i values and $P_{\tau_i}^{(l)}(\cdot)$values be intervals: $\chi_i = \left[\underline{\chi}_i; \overline{\chi}_i\right]$, $P_{\tau_i}^{(l)} = \left[\underline{P}_{\tau_i}^{(l)}; \overline{P}_{\tau_i}^{(l)}\right]$; let w be a non-negative monotonically decreasing function defined over interval [0;1] and let $l > 1$ be a real number:

Definition 11. The Weighted Fuzzy Expected Interval $WFEI_{gl}$ of order l with the attached weight function w of the compatibility function $\chi_{\widetilde{A}}$ with respect to fuzzy measure g is called the limit of the iteration process of the combinatorial interval extension :

$$s_n = \frac{\sum\limits_{i=1}^{k} \left[\underline{\chi}_i; \overline{\chi}_i\right] w\left(\left|\left[\underline{\chi}_i; \overline{\chi}_i\right] - s_{n-1}\right|\right) \left[\underline{P}_{\tau_i}^{(l)}\left(x_{\tau_i(1)}\right); \overline{P}_{\tau_i}^{(l)}\left(x_{\tau_i(1)}\right)\right]}{\sum\limits_{i=1}^{k} w\left(\left|\left[\underline{\chi}_i; \overline{\chi}_i\right] - s_{n-1}\right|\right) \left[\underline{P}_{\tau_i}^{(l)}\left(x_{\tau_i(1)}\right); \overline{P}_{\tau_i}^{(l)}\left(x_{\tau_i(1)}\right)\right]}, \qquad (6.14)$$

where $s_0 = FEV\left(\chi_{\widetilde{A}}\right)$. It is denoted as

$$WFEI_{gl} = WFEV_{gl}\left(\chi_{\widetilde{A}}, w\right).$$

It's clear that $WFEI_{gl}$ is an interval extension of the $WFEV_{gl}$ and we have the following propositions:

Proposition 5. If $FEV = FEI$, intervals of compatibility values χ_i and values of associated probabilities (or values of fuzzy measure g) $P_\tau(\cdot)$ are one point intervals, then

$$WFEI_{gl} = WFEV_{gl}.$$

Clearly, the proof is trivial.

Proposition 6. If $X = \{x_1, ..., x_k\}$ is the set of variational sampling $\sim \begin{pmatrix} (x_1, x_2, ..., x_k) \\ (n_1, n_2, ..., n_k) \end{pmatrix}$ and $g : 2^X \to [0; 1]$ is "sampling" fuzzy measure:

$$g = g_{sampling},$$

then the following Generalized Weighted Fuzzy Expected values coincide:

$$WFEV_{gl} = WFEV, WFEI_{gl} = WFEI.$$

Clearly, the proof is trivial.

Conclusion: It can be stated that when there are insufficient data on population groups, the process of fuzzy statistical estimation comprises two stages: The generalisation of the fuzzy weighted estimator follows from a small amount of information, which is formally constructed by interval analysis and creates entropy growth of information. But, on the other hand, the mobile FSK principle leads to an entropy decrease of information, which is condensed in generalised fuzzy statistics and in the new population MTV, which is called the Weighted Fuzzy Expected Intervals ($WFEI$ and $WFEI_{gl}$) and the Weighted Fuzzy Expected value with respect to fuzzy measure g ($WFEV_{gl}$).

References

1. Campos I., Bolanos C. *Representation of Fuzzy Measures Through Probabilities.* Fuzzy Sets and Systems 31 (1989) 23-36.

2. Dong W.M., Wong F.S. *Fuzzy Weighted Averages and Implemetation of the Extension Principle.* Fuzzy Sets and Systems 21(1987) 183-199.

3. Dubois D., Prade H. *Theory of Possibility - An Approach to Computerized Processing of Uncertainty.* Plenum Press. New York,1988.

4. Friedman M., Schneider M., Kandel A. *The use of Weighted Fuzzy Expected Value (WFEV) in Fuzzy Expert Systems.* Fuzzy Sets and Systems 31(1989) 37-45.

5. Kandel A. *Fuzzy Statistics and Forecast Evaluation.* IEEE Trans. on Systems, Man and Cybernetics, SMG-8, No 5 (1978), 396-401.

6. Kandel A. *on the Control and Evaluation of Uncertain Processes.* IEEE Trans. on Automatic Control, vol AC-25, No.6 (1980), 1182-1187.

7. Moore R.E. *Interval Analysis.* Englewood Cliffs. N.J.Prentice-Hall,1996.

8. Schneider M., Kandel A. *Properties of the Fuzzy Expected Value and the Fuzzy Expected Interval in Fuzzy Environment.* Fuzzy Sets and Systems 28 (1988) 55-68.

9. Schneider M., Friedman M., Kandel A. *On fuzzy reasoning in expert systems.* Proceeding of the 1987 International Symposium on Multiple-valued Logic, Boston, MA (1987).

10. Sirbiladze G. *Insufficient Data and Weighted Fuzzy Expected Interval of Population.* Proc. Tbilisi State University. (1999) to be published.

11. Sugeno M. *Theory of Fuzzy Integrals and Its Applications.* Ph. Doctoral Thesis. Tokyo Institute of Technology. (1974)