# CHAPTER I

# CHAPTER – I

# INSURANCE USES OF FUZZY LOGIC

The purposes of this chapter are twofold: first, to document the FL(Fuzzy Logic) technologies employed in insurance-related areas; and, second, to review the FL applications so as to document the unique characteristics of insurance as an application area.

This chapter contains a brief overview of insurance application areas. It is subdivided by the fuzzy techniques shown in Fig. 1.

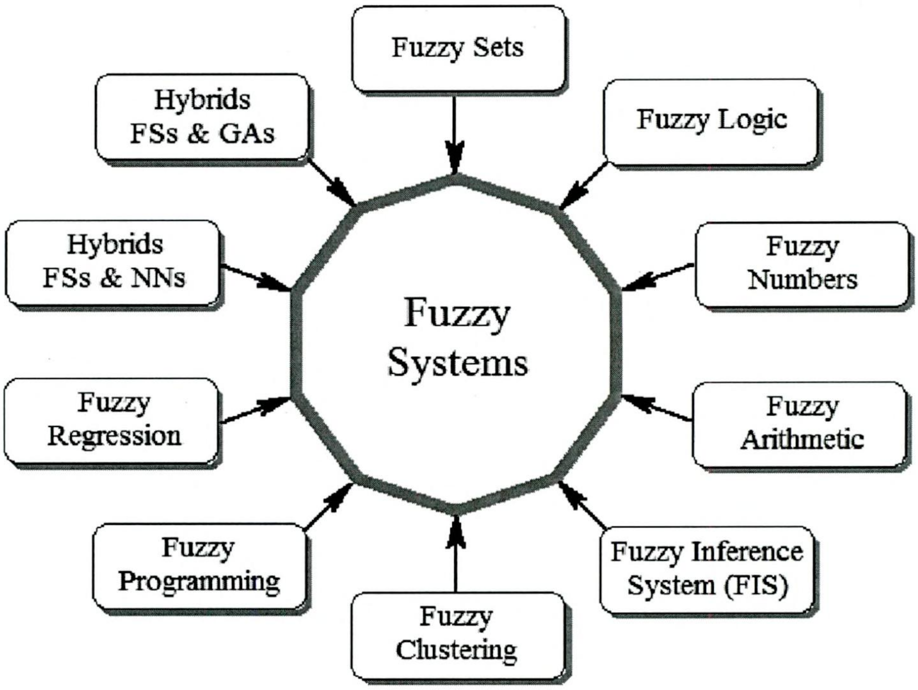


**Fig.1.** Fuzzy Logic

The topics covered include fuzzy set theory, fuzzy numbers, fuzzy arithmetic, fuzzy inference systems, fuzzy clustering, fuzzy programming, fuzzy regression, and soft computing. Each section begins with a brief description of the technique and is followed by a chronological review of the insurance applications of that technique.

# SECTION – 1.1

# INSURANCE APPLICATION AREAS

The major application areas of insurance include classification, underwriting, projected liabilities, ratemaking and pricing, and asset allocations and investments. In this section, we briefly describe each of these areas so that readers who are unfamiliar with the insurance field will have a context for the rest of the paper.

## 1.1.1. Classification

Classification is fundamental to insurance. On the one hand, classification is the prelude to the underwriting of potential coverage, while on the other hand, risks need to be properly classified and segregated for pricing purposes. Operationally, risk may be viewed from the perspective of the four classes of assets (physical, financial, human, intangible) and their size, type, and location.

## 1.1.2. Underwriting

Underwriting is the process of selection through which an insurer determines which of the risks offered to it should be accepted, and the conditions and amounts of the accepted risks. The goal of underwriting is to obtain a safe, yet profitable, distribution of risks. Operationally, underwriting determines the risk associated with an applicant and either assigns the appropriate rating class for an insurance policy or declines to offer a policy.

## 1.1.3. Projected Liabilities

In the context of this article, projected liabilities are future financial obligations that arise either because of a claim against and insurance company or a contractual benefit agreement between employers and their employees. The evaluation of projected liabilities is fundamental to the insurance and employee benefit industry, so it is not surprising that we are beginning to see SC technologies applied in this area.

### 1.1.4. Ratemaking and Pricing

Ratemaking and pricing refer to the process of establishing rates used in insurance or other risk transfer mechanisms. This process involves a number of considerations including marketing goals, competition and legal restrictions to the extent they affect the estimation of future costs associated with the transfer of risk. Such future costs include claims, claim settlement expenses, operational and administrative expenses, and the cost of capital.

### 1.1.5. Asset Allocation and Investments

The analysis of assets and investments is a major component in the management of an insurance enterprise. Of course, this is true of any financial intermediary, and many of the functions performed are uniform across financial companies. Thus, insurers are involved with market and individual stock price forecasting, the forecasting of currency futures, credit decision-making, forecasting direction and magnitude of changes in stock indexes, and so on.

## SECTION – 1.2

## LINGUISTIC VARIABLES AND FUZZY SET THEORY

Linguistic variables are the building blocks of FL. They may be defined ([58], [59]) as variables whose values are expressed as words or sentences. Risk capacity, for example, a common concern in insurance, may be viewed both as a numerical value ranging over the interval [0,100%], and a linguistic variable that can take on values like high, not very high, and so on. Each of these linguistic values may be interpreted as a label of a fuzzy subset of the universe of discourse X = [0,100%], whose base variable, x, is the generic numerical value risk capacity. Such a set, an example of which is shown in Fig. 2, is characterized by a membership function (MF), $\mu\ high(x)$ here, which assigns to each object a grade of membership ranging between zero and one.

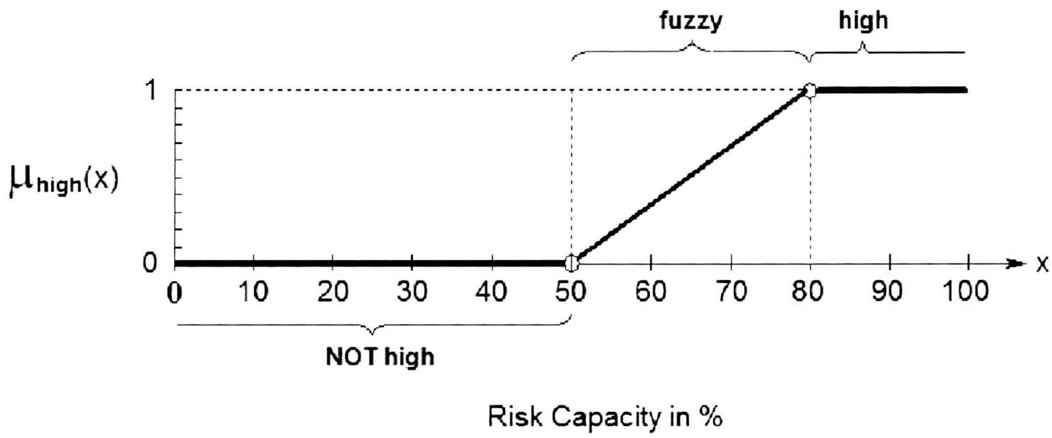


**Fig.2.** (Fuzzy) Set of Clients with High Risk Capacity

In this case, which represents the set of clients with a high risk capacity, individuals with a risk capacity of 50 percent, or less, are assigned a membership grade of zero and those with a risk capacity of 80 percent, or more, are assigned a grade of one. Between those risk capacities, (50%, 80%), the grade of membership is fuzzy.

In addition to the S-shaped MF depicted in Fig. 2, insurance applications also employ the triangular, trapezoidal, Gaussian, and generalized bell classes of MFs. As with other areas of application, fuzzy sets are implemented by extending many of the basic identities that hold for ordinary sets.

### 1.2.1. Applications

This subsection presents an overview of some insurance applications of linguistic variables and fuzzy set theory. The topics addressed include: earthquake insurance, optimal excess of loss retention in a reinsurance program, the selection of a “good” forecast, where goodness is defined using multiple criteria that may be vague or fuzzy.

An early study was by [7], who used pattern recognition and FL in the evaluation of seismic intensity and damage forecasting, and for the development of models to estimate earthquake insurance premium rates and insurance strategies. The influences on the performance of structures include quantifiable factors, which can be captured by probability models, and nonquantifiable factors, such as construction quality and architectural details, which are best formulated using fuzzy set models. For example, he defined the percentage of a building damaged by an earthquake by fuzzy terms such as medium, severe and total, and represented the membership functions of these terms as shown in Fig. 3.

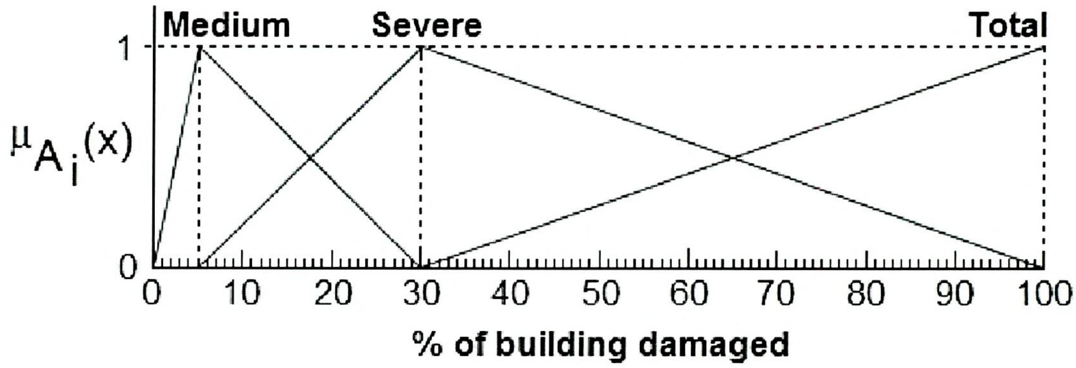


**Fig.3.** MFs of Building Damage

Two methods of identifying earthquake intensity were presented and compared. The first method was based on the theory of pattern recognition where a discriminative function was developed using Bayes' criterion and the second method applied FL.

Reference [40] envisioned the decision-making procedure in the selection of an optimal excess of loss retention in a reinsurance program as essentially a maximin technique, similar to the selection of an optimum strategy in noncooperative game theory. As an example, he considered four decision variables (two goals and two constraints) and their membership functions: probability of ruin, coefficient of variation, reinsurance premium as a percentage of cedent's premium income (Rel. Reins. Prem.) and deductible (retention) as a percentage of cedent's premium income (Rel. Deductible). The grades of membership for the decision variables (where the vertical lines cut the MFs) and their degree of applicability (DOA), or rule strength, may be represented as shown Fig. 4.

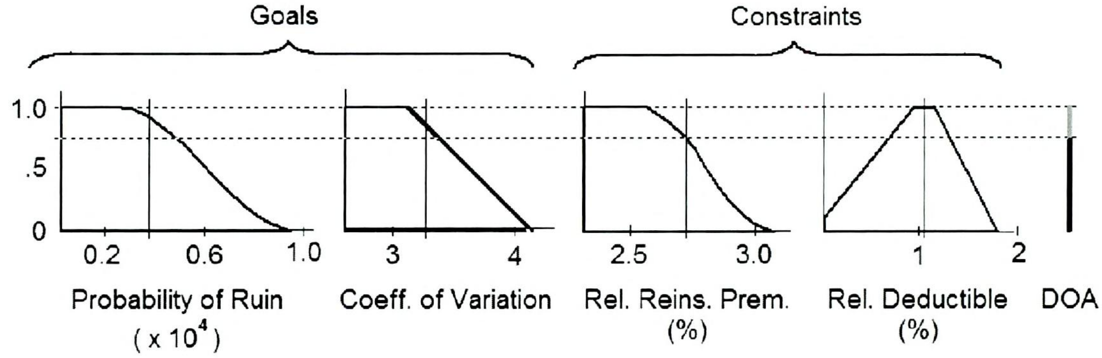


**Fig.4.** Retention Given Fuzzy Goals and Constraints

In the choice represented in the figure, the relative reinsurance premium has the minimum membership value and defines the degree of applicability for this particular excess of loss reinsurance program. The optimal program is the one with the highest degree of applicability.

Reference [18, p. 434] studied fuzzy trends in property-liability insurance claim costs as a follow-up to their assertion that "the actuarial approach to forecasting is rudimentary." The essence of the study was that they emphasized the selection of a "good" forecast, where goodness was defined using multiple criteria that may be vague or fuzzy, rather than a forecasting model. They began by calculating several possible trends using accepted statistical procedures and for each trend they determined the degree to which the estimate was good by intersecting the fuzzy goals of historical accuracy, unbiasedness and reasonableness.

The flavor of the article can be obtained by comparing the graphs in Fig. 5, which show the fuzzy membership values for 30 forecasts according to historical accuracy (goal 1), ordered from best to worst, and unbiasedness (goal 2), before intersection, graph (a) and after intersection, graph (b).

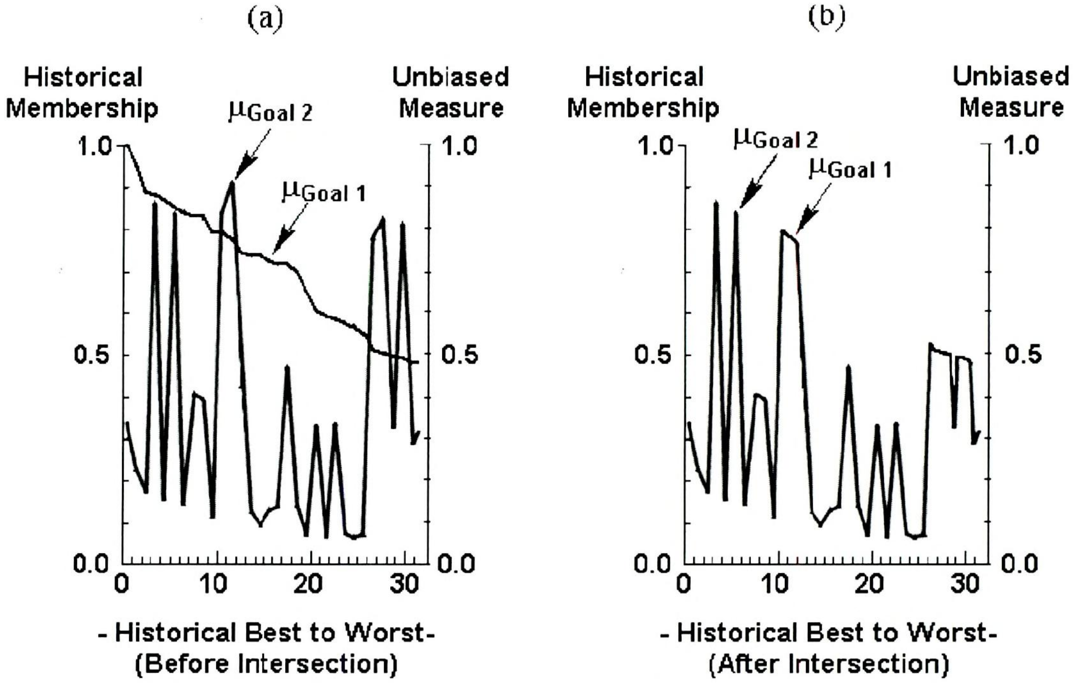


**Fig.5.** The Intersection of Historical Accuracy and Unbiasedness

They suggested that one may choose the trend that has the highest degree of goodness and proposed that a trend that accounts for all the trends can be calculated by forming a weighted average using the membership degrees as weights. They concluded that FL provides an effective method for combining statistical and judgmental criteria in insurance decision-making.

## SECTION – 1.3

## FUZZY NUMBERS AND FUZZY ARITHMETIC

### 1.3.1. Fuzzy Numbers

Fuzzy numbers are numbers that have fuzzy properties, examples of which are the notions of “around six percent” and “relatively high”. The general characteristic of a fuzzy number ([58] and [25]) often is represented as shown in Fig. 6, although any of the MF classes, such as Gaussian and generalized bell, can serve as a fuzzy number, depending on the situation.

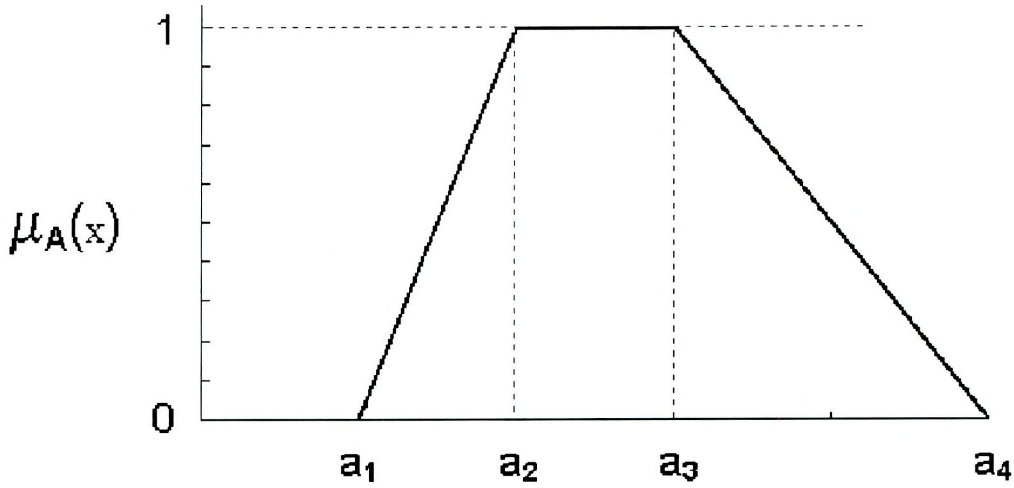


**Fig.6.** Flat Fuzzy Number

This shape of a fuzzy number is referred to as trapezoidal or "flat" and its MF often is denoted as $(a_1,a_2,a_3,a_4)$ or $(a_1/a_2, a_3/a_4)$; when $a_2$ is equal to $a_3$, we get the triangular fuzzy number. A fuzzy number is positive if $a_1 \geq 0$ and negative if $a_4 \leq 0$, and, as indicated, it is taken to be a convex fuzzy subset of the real line.

### 1.3.2. Fuzzy Arithmetic

As one would anticipate, fuzzy arithmetic can be applied to the fuzzy numbers. Using the extension principle ([58]), the nonfuzzy arithmetic operations can be extended to incorporate fuzzy sets and fuzzy numbers. Briefly, if $*$ is a binary operation such as addition (+), min ($\wedge$), or max ($\vee$), the fuzzy number $z$, defined by $z = x * y$, is given as a fuzzy set by

$$\mu_z(w) = \vee_{u,v}\, \mu_x(u) \wedge \mu_y(v)\,,\ u, v, w \in \Re \tag{1}$$

subject to the constraint that $w = u * v$, where $\mu_x$ , $\mu_y$, and $\mu_z$ denote the membership functions of $x$, $y$, and $z$, respectively, and $\vee_{u,v}$ denotes the supremum over $u$, $v$.

A simple application of the extension principle is the sum of the fuzzy numbers $A$ and $B$, denoted by $A \oplus B = C$, which has the membership function:

$\mu_c(z) = \max\{\min [\mu_A(x), \mu_B(y)] : x + y = z\}$ (2)

The general nature of the fuzzy arithmetic operations is depicted in Fig. 7 for $A = (-1, 1, 3)$ and $B = (1, 3, 5)$

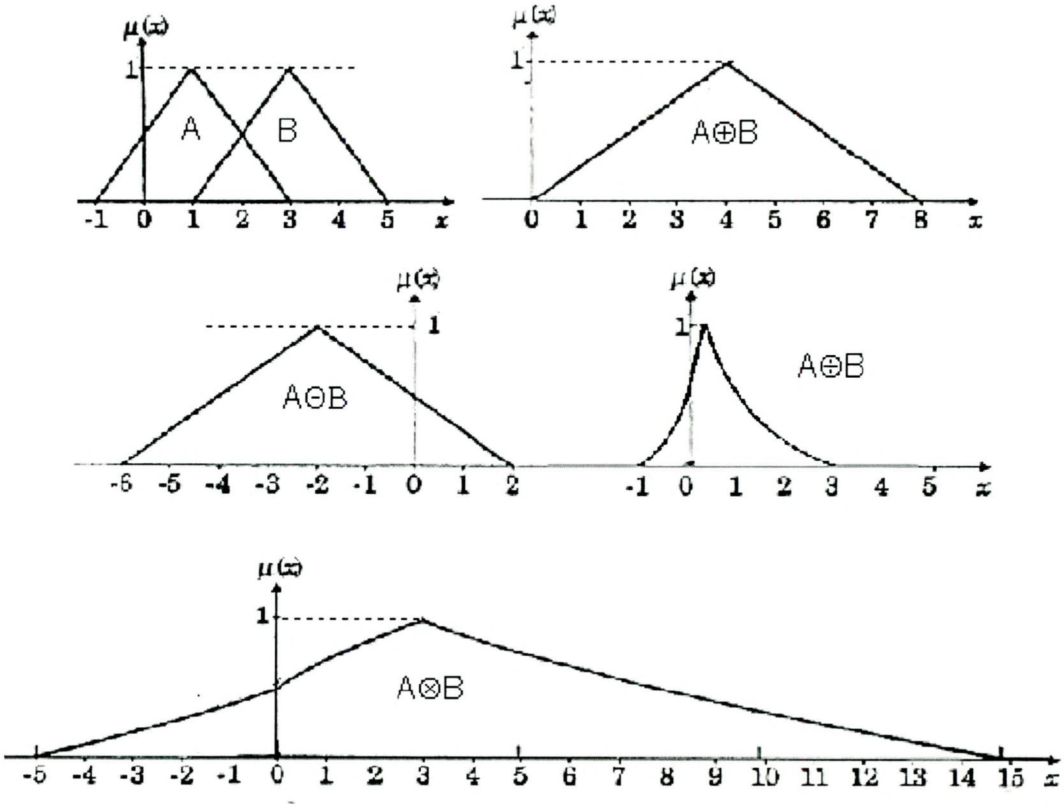


**Fig.7.** Fuzzy Arithmetic Operations

The first row shows the two membership functions $A$ and $B$ and their sum; the second row shows their difference and their ratio; and the third row shows their product.

## 1.3.3. Applications

This subsection presents an overview of insurance applications involving fuzzy arithmetic. The topics addressed include: the fuzzy future and present values of fuzzy cash amounts, using fuzzy interest rates, and both crisp and fuzzy periods; the computation of the fuzzy premium for a pure endowment policy.

Reference [9] appears to have been the first author to address the fuzzy time-value-of-money aspects of actuarial pricing, when he investigated the fuzzy future and present values of fuzzy cash amounts, using fuzzy interest rates, and both crisp and fuzzy periods. His approach, generally speaking, was based on the premise that "the arithmetic of fuzzy numbers is easily handled when x is a function of y." ([9, p. 258]) For a flat fuzzy number and straight line segments for $\mu A(x)$ on $[a_1, a_2]$ and $[a_3, a_4]$, this can be conceptualized as shown in Fig. 8

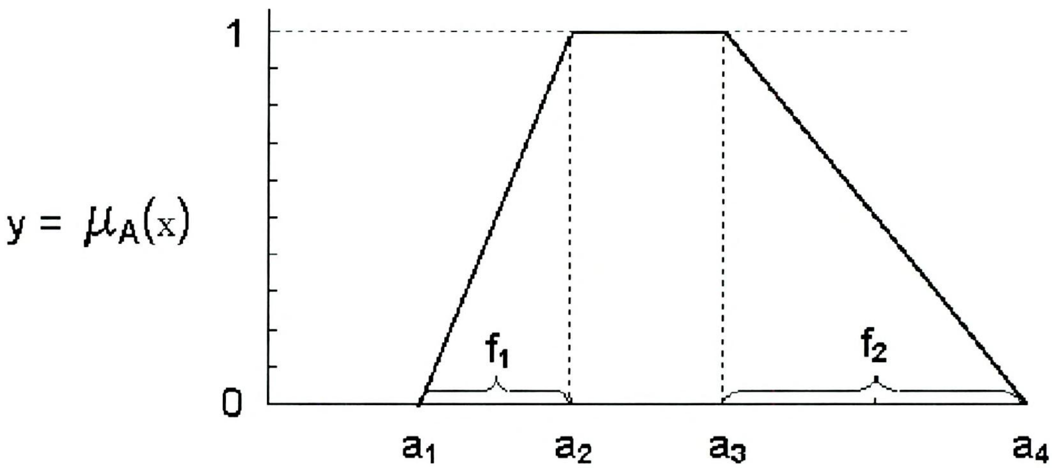


**Fig.8.** MF and Inverse MF

where $f_1(y|A) = a_1 + y(a_2 - a_1)$ and $f_2(y|A) = a_4 - y(a_4 - a_3)$. The points $a_j$ , $j = 1, 2, 3, 4$, and the functions $f_j(y|A)$, $j = 1, 2$, "A" a fuzzy number, which are inverse functions mapping the membership function onto the real line, characterize the fuzzy number.

If the investment is A and the interest rate per period is $i$, where both values are fuzzy numbers, he showed that the accumulated value ($S_n$), a fuzzy number, after n periods, a crisp number, is

$$S_n = A \otimes (1 \oplus i)^n \tag{3}$$

because, for positive fuzzy numbers, multiplication distributes over addition and is associative. It follows that the membership function for $S_n$ takes the form

$$\mu(x|S_n) = (s_{n1},\ f_{n1}(y|S_n)/s_{n2},\ s_{n3}/f_{n2}(y|S_n),\ s_{n4}) \tag{4}$$

where, for $j = 1, 2$,

$$f_{nj}(y|S_n) = f_j(y|A) \cdot (1 + f_j(y|i)) \tag{5}$$

and can be represented in a manner similar to Fig. 8, except that $a_j$ is replaced with $S_{nj}$.

Then, using the extension principle ([25]), he showed how to extend the analysis to include a fuzzy duration.

Buckley then went on to extend the literature to fuzzy discounted values and fuzzy annuities. In the case of positive discounted values, he showed that:

If $S > 0$ then $PV_2(S, n)$ exists; otherwise it may not, where:

$$PV_2(S, n) = A \text{ if } A \text{ is a fuzzy number and } A = S \otimes (1 \oplus i)^{-n} \tag{6}$$

The essence of his argument was that this function does not exist when using it leads to contradictions such as $a_2 < a_1$ or $a_4 < a_3$.

The inverse membership function of $PV_2(S, n)$ is:

$$f_j(y|A) = f_j(y|S) \cdot (1 + f_{3-j}(y|i))^{-n},\ j = 1, 2 \tag{7}$$

Both the accumulated value and the present value of fuzzy annuities were discussed.

Reference [40], using [9] as a model, discussed the computation of the fuzzy premium for a pure endowment policy using fuzzy arithmetic. Figure 9 is an adaptation of his representation of the computation.

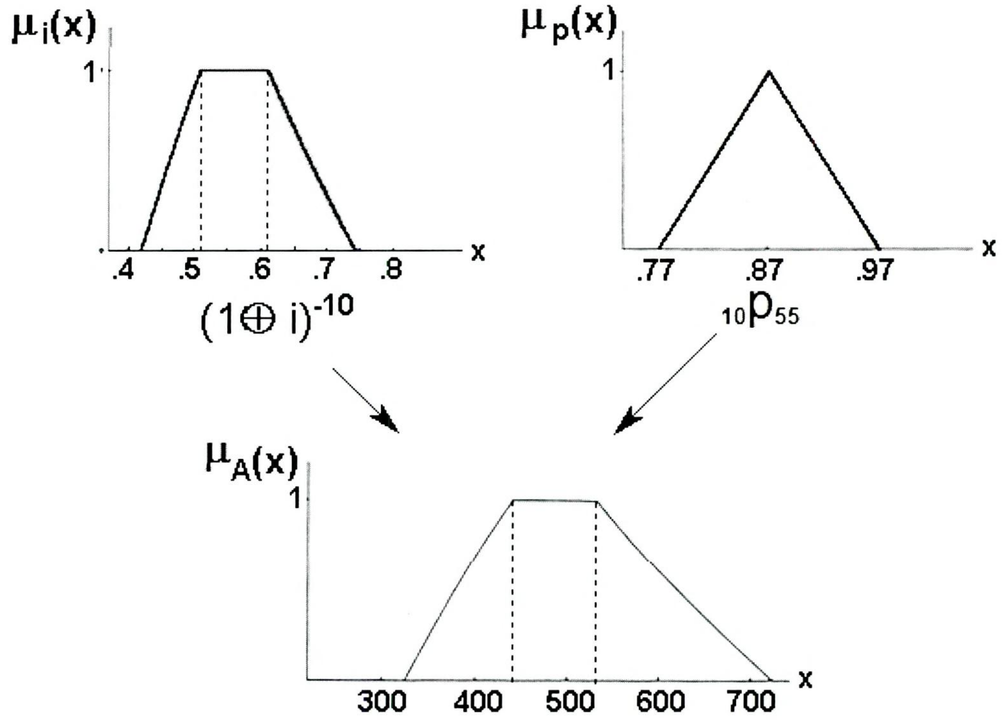


**Fig.9.** Fuzzy Present Value of a Pure Endowment

As indicated, the top left figure represents the MF of the discounted value after ten years at the fuzzy effective interest rate per annum of (.03, .05, .07, .09), while the top right figure represents the MF of ${}_{10}p_{55}$, the probability that a life aged 55 will survive to age 65. The figure on the bottom represents the MF for the present value of the pure endowment.

## SECTION – 1.4

## FUZZY INFERENCE SYSTEMS

The fuzzy inference system (FIS) is a popular methodology for implementing FL. FISs are also known as fuzzy rule based systems, fuzzy expert systems (FES), fuzzy models, fuzzy associative memories (FAM), or fuzzy logic controllers when used as controllers ([35, p. 73]), although not everyone agrees that all these terms are synonymous. Reference [5, p.77], for example, observes that a FIS based on IFTHEN rules is practically an expert system if the rules are developed from expert knowledge, but if the rules are based on common sense reasoning then the term expert system does not apply. The essence of a FIS can be represented as shown in Fig. 10.

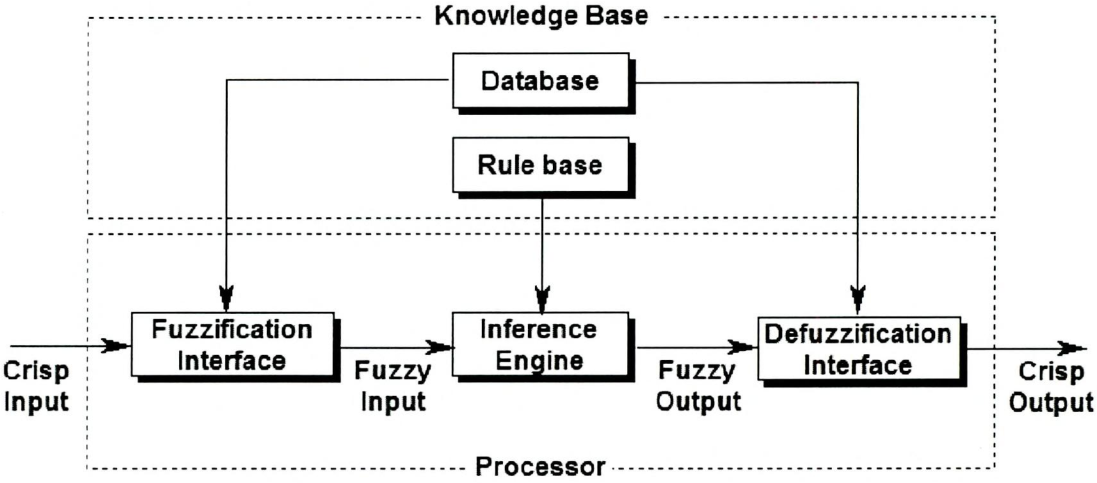


**Fig.10.** Fuzzy Inference System (FIS)

As indicated in the figure, the FIS can be envisioned as involving a knowledge base and a processing stage. The knowledge base provides MFs and fuzzy rules needed for the process. In the processing stage, numerical crisp variables are the input of the system. These variables are passed through a fuzzification stage where they are transformed to linguistic variables, which become the fuzzy input for the inference engine. This fuzzy input is transformed by the rules of the inference engine to fuzzy output. These linguistic results are then changed by a defuzzification stage into numerical values that become the output of the system.

The Mamdani FIS has been the most commonly mentioned FIS in the insurance literature, and most often the t-norm and the t-conorm are the min-operator and max-operator, respectively. Commonly, the centre of gravity (COG) approach is used for defuzzification.

### 1.4.1. Applications

This subsection presents an overview of insurance applications of FISs. In most instances, as indicated, an FES was used. The application areas include: life and health underwriting; classification; modeling the selection process in group health insurance.

As mentioned above, the first recognition that fuzzy systems could be applied to the problem of individual insurance underwriting was due to [21]. He recognized that underwriting was subjective and used a basic form of the FES to analyze the underwriting practice of a life insurance company.

Using what is now a common approach, he had underwriters evaluate 30 hypothetical life insurance applications and rank them on the basis of various attributes. He then used this information to create the five membership functions: technical aspects ($\mu_t$), health ($\mu_h$), profession ($\mu_p$), commercial ($\mu_c$), and other ($\mu_o$). Table 1 shows Dewit's conceptualization of the fuzzy set "technical aspects."

**Table 1.** Technical Aspects

| Description | Example | Fuzzy value |
|---|---|---|
| good | remunerative, good policy | 1.0 |
| moderate | unattractive policy provisions | 0.5 |
| bad | sum insured does not match wealth of insured | 0.2 |
| impossible | child inappropriately insured for large amount | 0.0 |

Next, by way of example, he combined these membership functions and an array of fuzzy set operations into a fuzzy expert underwriting system, using the formula:

$$W = \left(I(\mu t)\mu h \sqrt{\mu_p}\, \mu_o^2 \sqrt{2\min(0.5, \mu_c)}\right)^{[1-\max(0,\mu_c-0.5)]} \quad (8)$$

where intensification ($I(\mu_t)$) increases the grade of membership for membership functions above some value (often 0.5) and decreases it otherwise, concentration ($\mu_o^2$) reduces the grade of membership, and dilation ($\sqrt{\mu_p}$) increases the grade of membership. He then suggested hypothetical underwriting decision rules related to the values of W.

Reference [40] used a FES to provide a flexible definition of a preferred policyholder in life insurance. As a part of this effort, he extended the insurance underwriting literature in three ways: he used continuous membership functions; he extended the definition of intersection to include the bounded difference, Hamacher and Yager operators; and he showed how $\alpha$-cuts could be implemented to refine the decision rule for the minimum operator, where the $\alpha$-cuts is applied to each membership function, and the algebraic product, where the minimum acceptable product is equal to the $\alpha$-cut. Whereas [21] focused on technical and behavioral features, Lemaire focused on the preferred policyholder underwriting features of cholesterol, blood pressure, weight and smoker status, and their intersection.

An early classification study was [27], which discussed how measures of fuzziness can be used to classify life insurance risks. They envisioned a two-stage process. In the first stage, a risk was assigned a vector, whose cell values represented the degree to which the risk satisfies the preferred risk requirement associated with that cell. In the second stage, the overall degree of membership in the preferred risk category was computed. This could be done using the fuzzy intersection operator of [40] (see Fig. 4) or a fuzzy inference system. Measures of fuzziness were compared and discussed within the context of risk classification, both with respect to a fuzzy preferred risk whose fuzziness is minimized and the evaluation of a fuzzy set of preferred risks.

Following Lemaire's lead ([40]), [33] and [55] used FES to model the selection process in group health insurance. First single-plan underwriting was considered and then the study was extended to multiple-option plans. In the single-plan situation, Young focused on such fuzzy input features as change in the age/sex factor in the previous two years, change in the group size, proportion of employees selecting group coverage, proportion of premium for the employee and the dependent paid by the employer, claims as a proportion of total expected claims, the loss ratio, adjusted for employer size, and turnover rate. She completed the section with a discussion of a matrix of the interaction between the features (criteria) and their interpretation in the context of fuzzy intersection operators.

In the multiple-option case, the additional fuzzy features include single and family age factors, desired participation in each plan, age/sex factors, the difference in the cost of each plan, and the relative richness of each plan. The age factors depended on the possibility of participation, given access cost, the richness of the benefits, employee cost, marital status, and age. The underwriting decision in this case included the single-plan decision as a criterion.

## SECTION – 1.5

## FUZZY CLUSTERING

The foregoing fuzzy system allows us to convert and embed empirical qualitative knowledge into reasoning systems capable of performing approximate pattern matching and interpolation. However, these systems cannot adapt or learn because they are unable to extract knowledge from existing data. One approach for overcoming this limitation is to use fuzzy clustering. The essence of fuzzy clustering is that it produces reasonable centers for clusters of data, in the sense that the centers capture the essential feature of the cluster, and then groups data vectors around cluster centers that are reasonably close to them.

The fuzzy c-means algorithm ([6]) is a fuzzy clustering method referred to by a number of studies mentioned in this review. A flowchart of the algorithm is depicted in Fig. 11:

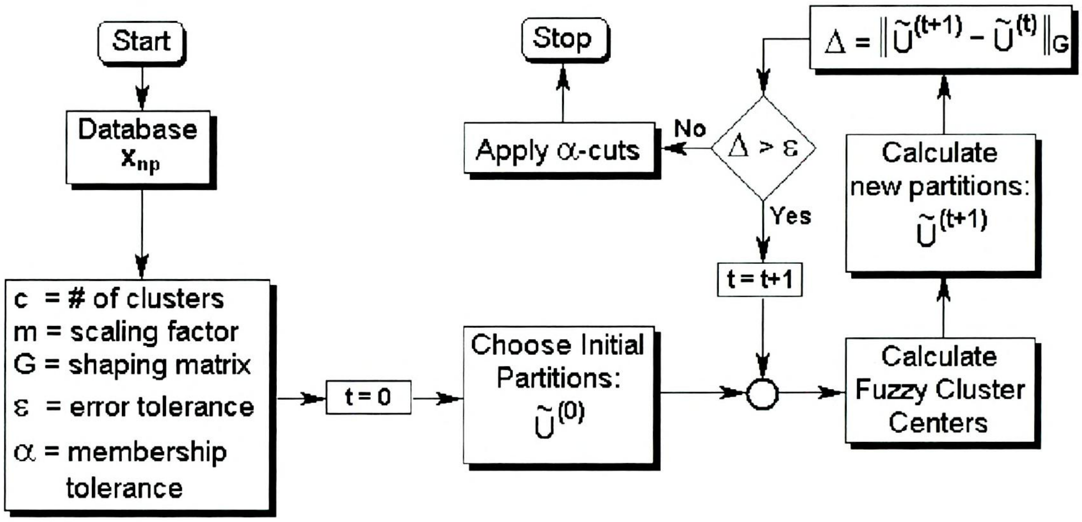


**Fig.11.** Flowchart of c-Means Algorithm

As indicated, the database consists of the $n \times p$ matrix, $x_{np}$, where $n$ indicates the number of patterns and $p$ denotes the number of features. The algorithm seeks to segregate these $n$ patterns into $c$, $2 \leq c \leq n - 1$, clusters, where the within clusters variances are minimized and the between clusters variances are maximized. To this end, the algorithm is initialized by resetting the counter, $t$, to zero, and choosing: $c$, the number of clusters; $m$, the exponential weight, which acts to reduce the influence of noise in the data because it limits the influence of small values of membership functions; $G$, a symmetric, positive-definite (all its principal minors have strictly positive determinants), $p \times p$ shaping matrix, which represents the relative importance of the elements of the data set and the correlation between them, examples of which are the identity and covariance matrixes; □ , the tolerance, which controls the stopping rule; and $\alpha$, the membership tolerance, which defines the relevant portion of the membership functions.

Given the database and the initialized values, the counter, $t$, is set to zero. The next step is to choose the initial partition (membership matrix), $\tilde{U}^{(0)}$, which may be based on a best guess or experience. Next, the fuzzy cluster centers are computed, which, in effect, are elements that capture the essential feature of the cluster. Using these fuzzy cluster centers, a new (updated) partition, $\tilde{U}^{(t+1)}$, is calculated. The partitions are compared using the matrix norm $\|\tilde{U}^{(t+1)} - \tilde{U}^{(t)}\|_G$ and if the difference exceeds □, the counter, $t$, is increased and the process continues. If the difference does not exceed □, the process stops. As part of this final step, $\alpha$-cuts are applied to clarify the results and make interpretation easier, that is, all membership function values less than $\alpha$ are set to zero and the function is renormalized.

### 1.5.1. Applications

This subsection presents an overview of insurance applications of the c-means algorithm. The application areas include: an alternate tool for estimating credibility; risk classification in both life and non-life insurance.

Reference [45] explored the use of fuzzy clustering methods as an alternate tool for estimating credibility. Given $\{x_{ij}, i = 1, \ldots, n, j = 1, \ldots, p\}$, a data set representing historical loss experience, and $y = \{y_j, j = 1, \ldots, p\}$, a data set representing the recent experience (risk characteristics and loss features), the essential idea is that one can use a clustering algorithm to assign the recent experience to fuzzy clusters in the data. Thus, if $\mu$ is the maximum membership degree of $y$ in a cluster, $Z = 1 - \mu$ could be used as the credibility measure of the experience provided by $y$, while $\mu$ gives the membership degree for the historical experience indicated by the cluster.

As an example they consider an insurer with historical experience in three large geographical areas extending its business to a fourth large area. The insurer can cluster new data from this fourth area into patterns from the other areas, and thereby derive a credibility rating for its loss experience in the new market. Using the c-means algorithm, the means and standard deviations as features, and two partitions ($c = 2$), they arrived at the credibility factor for the data of the fourth area.

Reference [46, Chap. 6] observed that lack of actuarially fair classification is economically equivalent to price discrimination in favor of high risk individuals and suggested "... a possible precaution against [discrimination] is to create classification methods with no assumptions, but rather methods which discover patterns used in classification." To his end, he was among the first to suggest the use of the c-means algorithm for classification in an insurance context.

By way of introducing the topic to actuaries, he discussed an insightful example involving the classification of four prospective insured's, two males and two females, into two clusters, based on the features height, gender, weight, and resting pulse. The two initial clusters were on the basis of gender. In a step-by-step fashion through three iterations, Ostaszewski developed a more efficient classification based on all the features.

Reference [19] and [20] extended the work of [46, Chap. 6] by showing how the c-means clustering algorithm could provide an alternative way to view risk and claims classification. Their focus was on applying fuzzy clustering to the two problems of grouping towns into auto rating territories and the classification of insurance claims according to their suspected level of fraud. Both studies were based on Massachusetts automobile insurance data.

The auto rating territories portion of the study involved 350 towns and the 10 Boston rated subdivisions, with the features bodily injury (BI) liability, personal injury protection (PIP), property damage liability (PDL), collision, comprehensive, and a sixth category comprising the five individual coverages combined. The parameters of the c-means algorithm were five coverage partitions ($c$ = 5), which was the number of categories in a previous territory assignment grouping, a scaling factor of 2 ($m$ = 2), a tolerance of 5 percent (□ = 0.05), and an $\alpha$-cut of 20 percent.

Figure 12 shows a representation of the impact of the clustering algorithm when applied to the auto rating territories of a subset of 12 towns ($x - axis$) and five clusters ($y - axis$). The subscripts "I" and "F" denote the initial and final clusters, respectively.

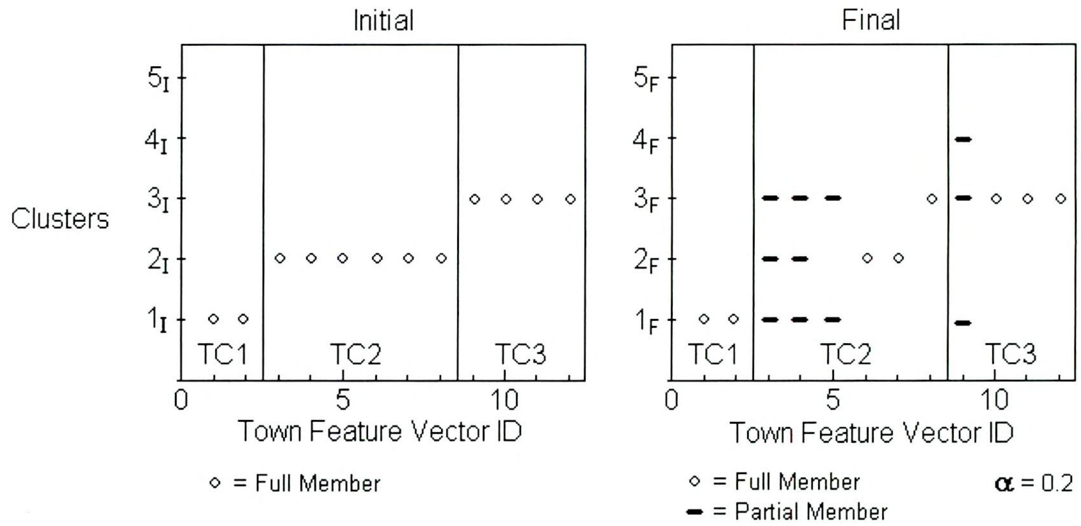


**Fig.12.** Town Clustering Using c-Means Algorithm

As indicated, in the left figure, the initial groups are crisp in the sense that the memberships of the territories are unique. In contrast, as a consequence of applying the c-means algorithm, the optimum classification resulted in some towns belonging to more than one cluster (risk class). Similar results were found for the entire database.

Their second study was based on an interesting use of information derived from experts. Beginning with 387 claims and two independent coders, 62 claims that were deemed fraudulent by either coder were identified. Then, starting with 127 claims (the 62 deemed fraudulent plus 65 from remaining 325), experienced claim managers and experienced investigators were each asked to rank each claim on a scale of 0 to 10. Their responses were grouped into the five initial clusters ($c = 5$): none (0), slight (1-3), moderate (4-6), strong (7-9), and certain (10). In this instance, the three features were the adjuster suspicion value, the investigator suspicion value, and a third category labeled the "fraud vote," which was equal to the number of reviewers who designated the claim as fraudulent. The results of their analysis supported the hypothesis that adjuster suspicion levels can serve well to screen suspicious claims. The authors concluded that fuzzy clustering is a valuable addition to the methods of risk and claim classification, but they did not conclude that it was the best way.

## SECTION – 1.6

## FUZZY PROGRAMMING

Many of the fuzzy logic studies in insurance involve decision making, and most of these studies rely on the framework established by [3]. The essential notion is that, given a non-fuzzy space of options, X, a fuzzy goal, G, and a fuzzy constraint, C, then G and C combine to form a decision, D, which is a fuzzy set resulting from the intersection of G and C. Assuming the goals and constraints enter into the expression for D in exactly the same way, a simple representation of the relationship between G, C and D is given in Fig. 13.

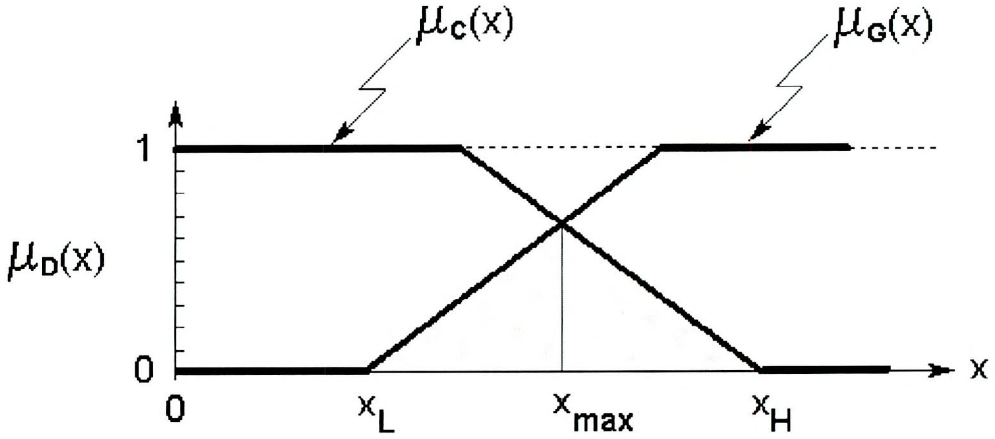


**Fig.13.** Decision Making

As indicated, the decision involves the fuzzy intersection of the goal and constraint MFs, and the set of possible options in the interval $x_L$ to $x_H$. If the optimal decision is the option with the highest degree of membership in the decision set, the crisp solution to this problem would be

$$X^* = \arg[\max_x \min\{\mu_G(x), \mu_C(x)\}]$$

In this section, we focus on the role of fuzzy linear programming (LP) in decision making. Like its crisp counterpart, fuzzy LP might involve finding an x such that ([66]: 289)

$$C = \sum_{ij} c_{ij}\, x_{ij} \lesseqgtr C_0$$

$$Z_i = \sum_j a_{ij}\, x_{ij} \lesseqgtr b_i$$

$$x_{ij} \geq 0$$

where $C_0$ is the aspiration level for the objective function, "~" over a symbol denotes the fuzzy version of that symbol, and the coefficients $a_{ij}$ , $b_i$, and $c_{ij}$ are not necessarily crisp numbers.

This fuzzy LP problem can be resolved by reformulating it as a crisp LP problem. The essence of one approach to doing this is depicted in Fig. 14.

As indicated, $z_i$ is a fuzzy number, whose membership function is zero for

$z_i \le b_i - \lambda_i$, one for $z_i \ge b_i$, and linearly increasing in the interval. Zimmermann refers to $\lambda$ as a tolerance interval. Using an $\alpha$-cut to provide a minimum acceptable satisfaction level, that is, $\mu(z_i) \ge \alpha$ is an acceptable constraint, we see from the diagram that an equivalent constraint is $z_i \ge b_i - \lambda_i + \lambda_i\alpha$. Similarly, $C \le C_0 + \lambda - \lambda\alpha$.

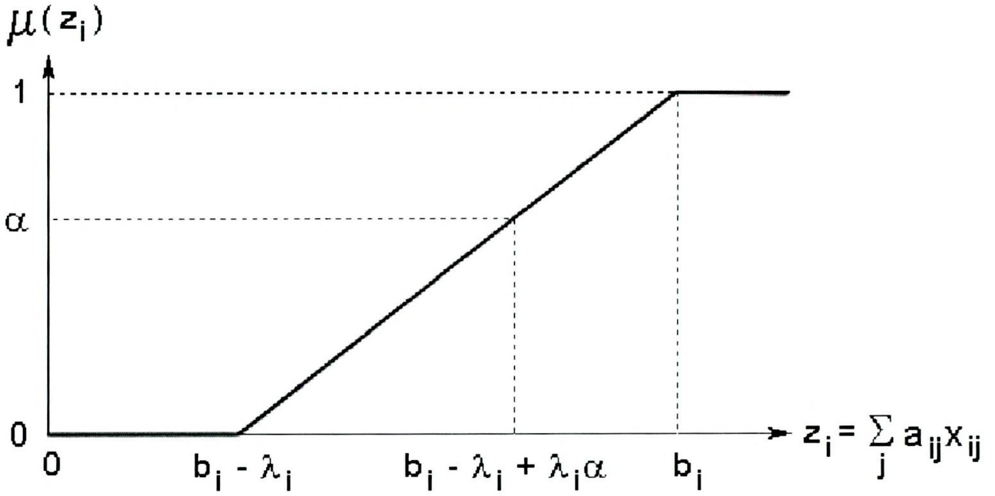


**Fig.14.** Equivalent Crisp Constraint

Thus, given the values of $\lambda$, the equivalent crisp programming problem becomes one of maximizing $\alpha$ subject to the equivalent constraints, that is:

$$Maximize : \alpha$$

$$Subject\ to : z_i - \lambda_i\alpha \ge b_i - \lambda_i;$$

$$C + \lambda\alpha \le C_o + \lambda;\ and$$

$$0 \le \alpha \le 1$$

### 1.6.1. Applications

A number of the foregoing articles used decision making, but, since they have already been reviewed, they will not be revisited here. Instead, we focus on three articles that explicitly incorporate linear programming. The topic addressed include optimal asset allocation, insurance pricing.

Reference [29] used a possibilistic linear programming method for optimal asset allocation based on simultaneously maximizing the portfolio return, minimizing the portfolio risk and maximizing the possibility of reaching higher returns. This was analogous to maximizing mean return, minimizing variance and maximizing skewness for a random rate of return.

The authors conceptualize the possibility distribution ($\pi_{\tilde{r}_i}$) of the imprecise rate of return of the $i-th$ asset of the portfolio as shown in Fig. 15(a), where $\tilde{r}_i = (r_i^p, r_i^m, r_i^o)$ and $r_i^p, r_i^m, r_i^o$ are the most pessimistic value, the most possible value, and the most optimistic value for the rate of return, respectively.

Then, as depicted in Fig. 15(b), taking the weighted averages of these values, they defined the imprecise rate of return for the entire portfolio as $\tilde{r} = (r^p, r^m, r^o)$, the portfolio risk as $(r^m - r^p)$, and the portfolio skewness as $(r^o - r^m)$. The authors then showed in a step-by-step fashion how the portfolio could be optimized using [61] fuzzy programming method. They concluded that their algorithm provides maximal flexibility for decision makers to effectively balance the portfolio's return and risk.

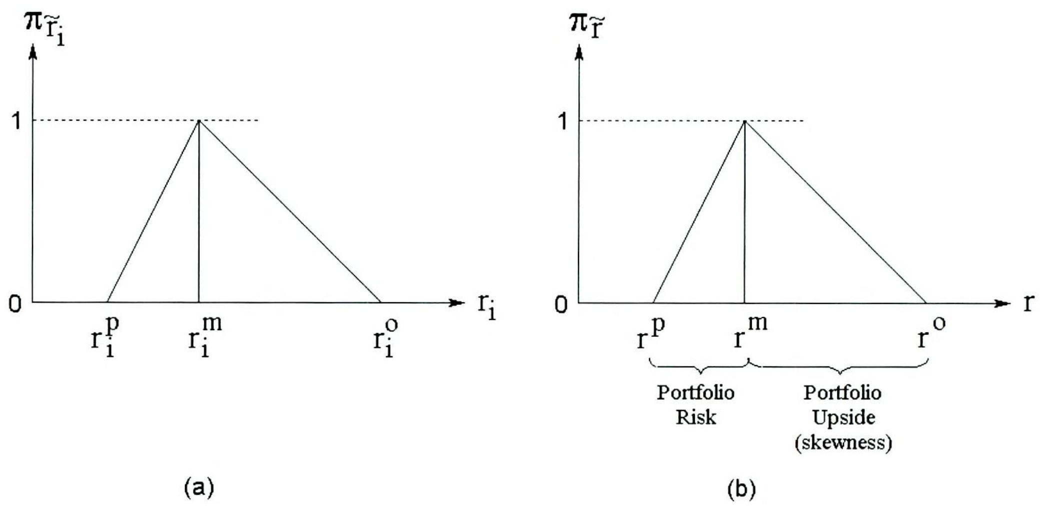


**Fig.15.** Possibility Distribution of Portfolio Return

Reference [12] investigated the use of fuzzy mathematical programming for insurance pricing decisions with respect to a bonus-malus rating system in automobile insurance. They used the max-min operator and followed Zimmermann's approach ([62], [63]), which led to an optimal solution of the form:

$$\mu_D(x^*) = max_x\, min_x\{\mu_O(x),\, \mu_{R_i}(x)\}\}, \; i = 1, \ldots, k \tag{9}$$

where $\mu_D$, $\mu_O$, and $\mu_R$ denote the membership function for the fuzzy set "decision D," the fuzzy objective function, and the fuzzy constraints, respectively. Their assumed objective was "attractive income from premiums" while the constraints involved the spread of policies among the risk classes, the weighted sum of the absolute variation of the insured's premium, and the deviation from perfect elasticity of the policyholder's payments with respect to their claim frequency. The system was tested on a large database of third-party personal liability claims of a Spanish insurer and they concluded that their fuzzy linear programming approach avoids unrealistic modeling and may reduce information costs. Reference [13] provides further commentary on the approach.

## SECTION – 1.7

## FUZZY REGRESSION

Two general approaches have been used to develop fuzzy regression models. The earlier approach, the possibilistic model of [51], took the general form

$$\tilde{Y} = \tilde{A}_0 + \tilde{A}_1 x_1 + \ldots + \tilde{A}_n x_n \tag{10}$$

where $\tilde{Y}$ is the fuzzy output, $\tilde{A}_j$ , $j = 1, 2, \ldots, n$, is a fuzzy coefficient, and $\mathbf{x} = (x_1, \ldots, x_n)$ is an n-dimensional non-fuzzy input vector. The fuzzy components were assumed to be triangular fuzzy numbers (TFNs). The methodology was to minimize the fuzziness of the model by minimizing the total spreads of its fuzzy coefficients, subject to including the data points of each sample within a specified feasible data interval.

The second approach, the least-squares model of [22], used distance measures to minimize the distance between the output of the model and the observed output. Essentially, he defined an $L^2$-metric $d(.,.)^2$ between two TFNs by [22, p. 143, Eq. (2)].

$$d\left(\langle m_1, l_1, r_1\rangle, \langle m_2, l_2, r_2\rangle\right)^2 =$$

$$(m_1 - m_2)^2 + ((m_1 - l_1) - (m_2 - l_2))^2 + ((m_1 + r_1) - (m_2 + r_2))^2 \quad (11)$$

As indicated, it measures the distance between two fuzzy numbers based on their modes, left spreads and right spreads.

### 1.7.1. Applications

This subsection presents an overview of some insurance applications involving fuzzy regression. The application includes life insurance sales as a function of price changes, the modeling of the term structure of interest rates.

Reference [38], in a study of life insurance sales as a response to price changes, sketched out the essence of a fuzzy regression model applied to the S-shaped response curve

$$S_i = \sigma_i P_i^{e_i} \prod_{j=1}^{n} P_j^{e_{ij}} f(i) \quad (12)$$

where $S_i$ is the sales response, $P_i$ is the price of the i-th product, $\sigma_i$ is a scaling coefficient, $e_i$ and $e_{ij}$ are direct and cross elasticities, respectively, and $f(i)$ accounts for factors such as advertising expenses and number of insurance agents. Their model incorporated the possibility measure and its counterpart, the necessity measure, and they focused on the case where no sample data is available, such as the new market economies of the former socialist countries. Unfortunately, the authors gave no detail about the application.

Reference [43] developed a fuzzy linear regression model to predict the relationship of known risk factors to the onset of occupational injury. Following [51], their model was constructed using four fuzzy sub-modules to represent the risk associated with task, anthropometrics, joint deviation, and personal factors. These modules were combined to produce an overall risk level and a final risk prediction model. The results indicate that fuzzy linear regression is a useful technique for addressing the uncertainty associated with the definition and modeling of occupational injury risk factors.

Reference [48] used possibilistic regression to analyze the Term Structure of Interest Rates (TSIR). Key components of their methodology included constructing a discount function from a linear combination of quadratic splines, the coefficients of which were assumed to be STFNs, and using the minimum and maximum negotiated price of fixed income assets to obtain the spreads of the dependent variable observations. Given the fuzzy discount functions, the authors provide TFN approximations for the spot rates and forward rates, and then go on to show how to use the discount functions to obtain the net single premiums for some basic life insurance contracts.

Reference [48] also used possibilistic regression to estimate incurred claims that have not yet been reported to the insurer, IBNR claims. Their used a common claims run-off model, but instead of OLS, they used the fuzzy linear relation

$$\tilde{Z}_{i+1,j} = \tilde{b}_i + \tilde{c}_i Z_{i,j} \tag{13}$$

where $Z_{i,j}$ is the accumulated incurred losses of accident year $j$ at the end of development year $i$ and $\tilde{b}_i$ and $\tilde{c}_i$ are STFNs. $\tilde{Z}_{n,j}$ is not a STFN because it is obtained by iteration, but the authors give a reasonable STFN approximation of it. The final phase of their model provides a FN for the IBNR reserve for the $j-th$ year of occurrence, which is summed to produce the whole IBNR reserve.

# SECTION – 1.8

# SOFT COMPUTING

Most of the previously discussed studies focused on FL to the exclusion of other technologies. While their approach has been productive, it may have been sub-optimal, in the sense that studies may have been constrained by the limitations of FL, and opportunities may have been missed to take advantage of potential synergies afforded by other technologies.

This notion was embodied in the concept of soft computing (SC), which was introduced by [60]. He envisioned SC as being "concerned with modes of computing in which precision is traded for tractability, robustness and ease of implementation." For the most part, SC encompasses the technologies of fuzzy logic, genetic algorithms (GAs), and neural networks (NNs), and it has emerged as an effective tool for dealing with control, modelling, and decision problems in complex systems. In this context, FL is used to deal with imprecision and uncertainty, GAs are used for search and optimization, and NNs are used for learning and curve fitting. In spite of these dichotomies, there are natural synergies between these technologies, the technical aspects of which are discussed in [49].

## 1.8.1. Applications

This section provides a brief overview of a few representative insurance-related articles that have merged FL with either GAs or NNs. The application area considered is classification and involves four representative SC articles in insurance, two on the property-casualty side and two on the life-health side.

Our first example of a SC approach is the study of [54], which proposed it as an auto insurance claim processing system for Korea. In Korea, given personal and/or property damage in a car accident, the compensation rate depends on comparative negligence, which is assigned using responsibility rates. The authors first describe the expert knowledge structure and the claims processing system.

They then explain in general terms how they determined the responsibility rate, and hence the compensation rate, using a fuzzy database, a rule based system, and a feed-forward NN learning mechanism, and the problems associated with implementing their system.

Reference [17] reported on a SC-based fraud and abuse detection system for managed healthcare. The essence of the system was that it detected "anomalous" behavior by comparing an individual medical provider to a peer group. The preparation of the system involved three steps: identify the proper peer population, identify behavior patterns, and analyze behavior pattern properties. The peer population was envisioned as a three-dimensional space composed of organization type, geographic region, and organization size.

The behavior patterns were developed using the experience of a fraud-detection department, an unsupervised NN that learnt the relationships inherent in the claim data, and a supervised approach that automatically generate a fuzzy model from a knowledge of the decision variables. Finally, the behavior pattern properties were analyzed using the statistical measures mean, variance, standard deviation, mean absolute deviation, Kolmogorov-Smirnov (KS) test, skewness, and kurtosis.

The discovery properties of the fuzzy model were based on three static and one time varying criteria metrics. The static metrics were the insurer's exposure to fraudulent behavior, as measured by total claim dollars, the degree of variance from the center of the peer population for each behavior pattern, which was referred to as the population compatibility number, and the number of behaviors that are significantly at variance. The time varying metric was the change in the behavior population dynamics over time. Given the prepared system and the discovery properties, the distribution of data points for the behavior patterns of any individual provider within this population could be computed and compared with all the providers of a similar type, a similar organization size, and within the same geographic area. Thus, the fuzzy system-based fraud and abuse detection system identifies a provider that has significant variance from the peer population.

Cox concluded that the system was capable of detecting anomalous behaviors equal to or better than the best fraud-detection departments.

Reference [47] used GA-constructed FISs to automatically produce diagnostic systems for breast cancer diagnosis. The Pittsburgh-style of GAs was used to generate the database and rulebase for the FISs, based on a data furnished by specialists, which contained 444 benign cases and 239 malignant cases, which had been evaluated based on 9 features. They claimed to have obtained the best classification performance to date for breast cancer diagnosis and, because their final systems involve just a few simple rules, high human-interpretability.

Reference [4] used an evolutionary-fuzzy approach to investigate suspicious home insurance claims, where genetic programming was employed to evolve FL rules that classified claims into "suspicious" and "non-suspicious" classes. Notable features of his methodology were that it used clustering to develop membership functions and committee decisions to identify the best-evolved rules. With respect to the former, the features of the claims were clustered into low, medium, and high groups, and the minimum and maximum value in each cluster was used to define the domains of the membership functions. The committee decisions were based on different versions of the system that were run in parallel on the same data set and weighted for intelligibility, which was defined as inversely proportional to the number of rules, and accuracy. Bentley reported that the results of his model when applied to actual data agreed with the results of previous analysis.