

UPDATING PARAMETERS WITH FUZZY ENTROPIES

By Tai-Yan Kam¹ and Colin B. Brown,² M. ASCE

ABSTRACT: The uncertainty of a continuous variable is estimated by the combination of the objective entropy, which is obtained from the statistics of a countable population, and the fuzzy entropy, which is a measure of imprecision estimated from an engineer's experience and knowledge. This is illustrated by using subjective, imprecise, fuzzy information to update the objective parameters of an initial truncated normal distribution by means of combining Shannon's objective information entropy and fuzzy entropies. The entropies are treated as information measures of uncertainty of both the objective and subjective kinds. Three fuzzy entropy functions are presented and their effects on the altered parameters illustrated.

INTRODUCTION

In civil engineering, variables such as loads, resistance of structural members, strength of materials, and annual runoff are modeled as continuous variables. This involves constructing probability distributions from the limited countable information available. For instance, the lifetime maximum live load is modeled as a gamma variate. In many practical applications the true parameters of a probability distribution are different from those obtained from statistics. The factors affecting the distribution usually cannot be quantified in an objective way, but rather, they can be estimated through an engineer's experience and knowledge in an imprecise way. The updating of the parameters of the initial distribution may be made by Bayes' principle if the appropriate conditional test information is available. However, when subjective opinion is the updating information the the fuzzy set theory of Zadeh (5) can be used to describe the imprecision. A measure of the imprecision is the fuzzy entropy: DeLuca and Termini (2) have defined one such nonprobabilistic entropy within the context of fuzzy set theory. Brown (1) has shown that fuzzy probabilities for discrete variables can be evaluated from the combination of Shannon's objective entropy and the DeLuca-Termini fuzzy entropy.

In this paper both symmetric and monotonically increasing fuzzy entropies are used to update the parameters of a truncated normal distribution in the light of subjective information. The intention is to study the entropy relationships between fuzzy sets constructed from verbal statements and the modified truncated normal distribution parameters. The paper begins by constructing Shannon's objective entropy for the truncated normal distribution which in itself may be considered as best describing the countable information. This distribution may result from a constrained program based on Jaynes maximum entropy principle (1) or from a statistical confidence test. The subjective information based on

¹Assoc. Prof., Dept. of Civ. Engrg., National Chiao Tung Univ., Taiwan.

²Prof., Civ. Engrg., Univ. of Washington, Seattle, Wash.

Note.—Discussion open until May 1, 1984. To extend the closing date one month, a written request must be filed with the ASCE Manager of Technical and Professional Publications. The manuscript for this paper was submitted for review and possible publication on September 15, 1983. This paper is part of the *Journal of Engineering Mechanics*, Vol. 109, No. 6, December, 1983. ©ASCE, ISSN 0733-9399/83/0006-1334/\$01.00. Paper No. 18422.

an engineer's experience and knowledge modifies this objective information. The entropy need not necessarily increase because no state change has occurred. The new, subjective information merely completes the picture of uncertainty which is measured by the combined (subjective and objective) entropy of the system. The subjective entropy may increase or decrease the uncertainty of information on a system state and therefore either increase or decrease the combined entropy. The subjective information, as displayed through fuzzy entropies, is reviewed, followed by the combination of objective and subjective entropies to provide a complete uncertainty measure for the system. The rest of the paper develops methods for altering the objective truncated normal distribution parameters when this combined entropy is known. Examples are provided and the results discussed.

Objective Entropy.—The objective probability distribution of a random variable is constructed from statistical data where the measure of uncertainty can be expressed by Shannon's entropy function. Consider a random variable X which has a probability density function $p(x)$. According to Shannon's principle (4), the uncertainty measure of the variable is

$$H = -\int p(x) \ln p(x) dx \dots\dots\dots (1)$$

in which H is defined as the objective entropy. If X is a truncated normal variate with given standard deviation, σ , and expected value, m , the probability density function is

$$p(x) = \begin{cases} 0 & x < a \\ \frac{g(x)}{F(b) - F(a)} & a \leq x \leq b \dots\dots\dots (2) \\ 0 & x > b \end{cases}$$

in which a, b = the lower and upper limits of x , respectively

$$g(y) = \frac{1}{\sqrt{2\pi} \sigma} \exp \left[-\frac{1}{2} \left(\frac{y - m}{\sigma} \right)^2 \right]; \quad -\infty \leq y \leq \infty$$

and $F(u) = \int_{-\infty}^u p(y) dy$

The corresponding objective entropy is then

$$H = \frac{1}{F(b) - F(a)} \left[-\ln \frac{1}{F(b) - F(a)} + \ln \sqrt{2\pi e} \sigma \right] \dots\dots\dots (3)$$

Fuzzy Entropy.—Information extracted from an engineer's experience and knowledge is defined as subjective information which cannot be expressed in a statistical sense. It has been shown that the subjective information can be expressed in terms of fuzzy supports in the fuzzy set theory, and the measure of imprecision can be made by a nonprobabilistic entropy termed the fuzzy entropy.

Consider a variable, Y , which can only take on y_1, y_2, \dots, y_n in an imprecise manner with fuzzy support, $f(y_i)$, for membership in y_i , then

the average imprecision associated with such support of Y being a member of y_i is the fuzzy entropy $G[f(y_i)]$. The desiderata proposed by DeLuca and Termini are that for each $f(y_i)$ there exists G_i such that

1. $G_i = 0$ if and only if $f = 0$ or $f = 1$.
2. G_i must be a maximum if and only if $f = 1/2$.
3. $G_i \geq \bar{G}_i$ where $\bar{f} \geq f$ for $f \geq 1/2$, and $\bar{f} \leq f$ for $f \leq 1/2$.

An expression for G_i meeting these is

$$G_i = -f(y_i) \ln f(y_i) - [1 - f(y_i)] \ln [1 - f(y_i)] \dots \dots \dots (4)$$

and the fuzzy entropy for the n cases of $f(y_i)$ is

$$G = \sum_{i=1}^n G_i \dots \dots \dots (5)$$

The entropy function G_i of Eq. 4 possesses a symmetry with respect to $f = 1/2$ which has been analyzed by Munro (3). In particular, Munro points out that the symmetry of the entropy ensures that information for $f(y_i) < 0.5$ is meaningless inasmuch as it should be reducing the support for the objective measures; whereas it provides symmetric support about $f(y_i) = 0.5$. The symmetry may be removed with the following desiderata:

1. $G_i = 0$ if and only if $f = 1$.
2. G_i must be a maximum if and only if $f = 0$.
3. $G_i \geq \bar{G}_i$ for $\bar{f} \geq f$.

These lead to many expressions for G_i . Two possible expressions are

$$G_i = -\frac{2}{2 - f_i} \ln \frac{2}{2 - f_i} \dots \dots \dots (6)$$

$$\text{and } G_i = (2 - f_i) \ln (2 - f_i) \dots \dots \dots (7)$$

Substituting either Eq. 6 or Eq. 7 into Eq. 5 will give the fuzzy entropy.

Combined Entropy.—The combination of objective entropy and fuzzy entropy results in a total measure of uncertainty. If E is defined as the total entropy of the variable X , then from Eqs. 1 and 5, E can be expressed as

$$E = H + KG \dots \dots \dots (8)$$

in which $K =$ a multiplier.

The inclusion of the fuzzy entropy with the objective entropy will alter the probability of the underlying variable. Expressed in terms of the altered probability density function, $\bar{p}(x)$, the combined entropy is

$$E = -\int \bar{p}(x) \ln \bar{p}(x) dx \dots \dots \dots (9)$$

If the probabilistic model of the variable, before and after the inclusion of fuzzy entropy, remains the same, then the altered probability distribution of the truncated normal variate expressed in terms of the new parameters is

$$\tilde{p}(x) = \begin{cases} 0 & x < a \\ \frac{\tilde{g}(x)}{\tilde{F}(b) - \tilde{F}(a)} & a \leq x \leq b \dots\dots\dots (10) \\ 0 & x > b \end{cases}$$

in which $g = (1/\sqrt{2\Pi}\bar{\sigma}) \exp [-1/2 (x - \bar{m}/\bar{\sigma})^2]$; $\tilde{F}(u) = \int_{-\infty}^u \tilde{g}(y)dy$; and $\bar{\sigma}$, \bar{m} = the altered standard deviation and expected value, respectively. From Eqs. 9 and 10, the total combined entropy of the truncated normal variate is

$$E = \frac{1}{\tilde{F}(b) - \tilde{F}(a)} \left[-\ln \frac{1}{\tilde{F}(b) - \tilde{F}(a)} + \ln \sqrt{2\Pi e} \bar{\sigma} \right] \dots\dots\dots (11)$$

Determination of Altered Parameters.—The two altered parameters of the truncated normal distribution can be estimated through the construction of two entropy equations which are derived from the global and local uncertainty analyses as described in the following:

Analysis of Global Uncertainty.—In this analysis, subjective opinions concerning the standard deviation of the variable is sought. A fuzzy set, A , expressing this subjective information in terms of a membership support in the entropy of the variable is defined as

$$A = f|E \dots\dots\dots (12)$$

The support, f , in Eq. 12 expresses the fuzzy support for the subjective relevance of the objective statistics. If these statistics are considered to be true then the support is subjectively set to be 1. The objective entropy is then the same as the combined entropy and $G = 0$. On the other hand, if the statistics are considered to be irrelevant to the real situation then the variable is assumed to have a uniform distribution and the support is 1/2 if the entropy function of Eq. 4 is used or 0 if Eqs. 6 or 7 are used.

If the symmetric fuzzy entropy function of Eq. 4 is used in the analysis, the combined entropy of Eq. 8 becomes

$$E = H + K[-f \ln f - (1 - f) \ln (1 - f)] \dots\dots\dots (13)$$

and the criterion for determining the multiplier K is $E = -\ln (1/b - a)$; $f = 1/2$.

$$\text{Then } K = \frac{\ln \frac{1}{b-a} + H}{\ln \frac{1}{2}} \dots\dots\dots (14)$$

In view of Eq. 3, Eq. 13 becomes

$$E = \frac{1}{F(b) - F(a)} \left[-\ln \frac{1}{F(b) - F(a)} + \ln \sqrt{2\Pi e} \sigma \right] + K[-f \ln f - (1 - f) \ln (1 - f)] \dots\dots\dots (15a)$$

in which, from Eqs. 3 and 14

$$K = \frac{\ln \frac{1}{b-a} + \frac{1}{F(b) - F(a)} \left[-\ln \frac{1}{F(b) - F(a)} + \ln \sqrt{2\pi e} \sigma \right]}{\ln \frac{1}{2}} \dots\dots (15b)$$

With given values of σ and m , the altered standard deviation, $\bar{\sigma}$, can be determined from Eqs. 11 and 15 for any fuzzy support provided that the expected value of the distribution remains unaltered.

Example 1. Let $\sigma = 500$, $m = 5,000$, $a = 0$, and $b = 10,000$. If the expected value is assumed to remain constant, then the altered standard deviation, $\bar{\sigma}$, is evaluated from Eqs. 11 and 15. The ratio $\bar{\sigma}/\sigma$ for different values of f is represented by curve A in Fig. 1.

When the nonsymmetric fuzzy entropy function of Eq. 6 is used in Eq. 13, then the criterion for determining K in Eq. 13 becomes

$$E = -\ln \frac{1}{b-a}; \quad f = 0 \dots\dots\dots (16)$$

which gives
$$K = \frac{\ln \frac{1}{b-a} + H}{\frac{1}{2} \ln \frac{1}{2}} \dots\dots\dots (17)$$

Thus the combined entropy can be expressed as

$$E = \frac{1}{F(b) - F(a)} \left[-\ln \frac{1}{F(b) - F(a)} + \ln \sqrt{2\pi e} \sigma \right] + K \left(-\frac{1}{2-f} \ln \frac{1}{2-f} \right) \dots\dots\dots (18)$$

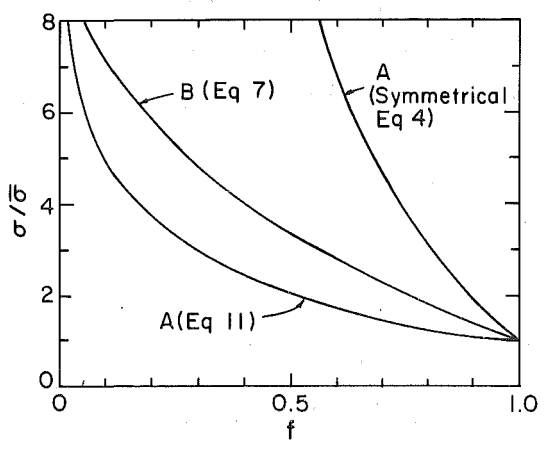


FIG. 1.—Change in $\bar{\sigma}/\sigma$ with Fuzzy Support, f for Different Fuzzy Entropy Definitions

Equating Eqs. 11 and 18 provides a relation between $\bar{\sigma}/\sigma$ and f which appears as curve B in Fig. 1 for the values of Example 1.

When the nonsymmetric fuzzy entropy function of Eq. 7 is used together with the criterion of Eq. 16, then

$$K = \frac{-\ln \frac{1}{b-a} - H}{2 \ln 2} \dots \dots \dots (19)$$

and the total or combined entropy is

$$E = \frac{1}{F(b) - F(a)} \left[-\ln \frac{1}{F(b) - F(a)} + \ln \sqrt{2\pi e} \sigma \right] + K[(2-f) \ln(2-f)] \dots \dots \dots (20)$$

Equating Eqs. 11 and 20 provides a third relation between $\bar{\sigma}/\sigma$ and f shown as C in Fig. 1.

Analysis of Local Uncertainty.—In most practical cases, both the standard deviation and expected value of a distribution are different from those obtained from objective statistics. Therefore, in order to determine both altered parameters, additional subjective information concerning the distribution of the variable is required.

From the objective probability distribution of Eq. 2, the probability of X being greater than $m - k\sigma$ is

$$P_1 = 1 - \frac{1}{F(b) - F(a)} \int_a^{m-k\sigma} p(x) dx \dots \dots \dots (21)$$

and of X being less than $m - k\sigma$ is

$$P_2 = 1 - P_1 \dots \dots \dots (22)$$

The objective entropy of these states, $X \geq m - k\sigma$ and $X < m - k\sigma$, is

$$H_L = P_1 \ln P_1 - (1 - P_1) \ln (1 - P_1) \dots \dots \dots (23)$$

Subjective opinion concerning the precision of these objective probabilities is then sought, and a measure of imprecision is introduced by a fuzzy entropy

$$G_L = P_1 G_1 + (1 - P_1) G_2 \dots \dots \dots (24)$$

If the symmetric fuzzy entropy function of Eq. 4 is adopted, then Eq. 23 becomes

$$G_L = P_1 [-f_1 \ln f_1 - (1 - f_1) \ln (1 - f_1)] + (1 - P_1) [-f_2 \ln f_2 - (1 - f_2) \ln (1 - f_2)] \dots \dots \dots (25)$$

and, the combined entropy, E_L , is

$$E_L = H_L + K_L G_L \dots \dots \dots (26)$$

The multiplier K_L is determined from the criterion (1) that when all $f_i = 1/2$ then $\bar{P}_i = 1/2$. Then

$$K_L = \frac{\ln \frac{1}{2} + H}{\ln \frac{1}{2}} \dots \dots \dots (27)$$

For the total or combined entropy of Eq. 26, the altered probabilities \tilde{p}_i are given in

$$E_L = -\tilde{P}_1 \ln \tilde{P}_1 - (1 - \tilde{P}_1) \ln (1 - \tilde{P}_1) \dots \dots \dots (28)$$

Therefore, \tilde{P}_1 can be determined from Eqs. 23–28.

For the altered distribution, Eq. 21 becomes

$$\tilde{P}_1 = 1 - \frac{1}{\tilde{F}(b) - \tilde{F}(a)} \int_a^{m-k\sigma} \tilde{p}(x) dx$$

$$\text{in which } \tilde{p}(x) = \frac{1}{\sqrt{2\Pi} \tilde{\sigma}} e^{-[1/2(x-\tilde{m})/\tilde{\sigma}]^2} \dots \dots \dots (29)$$

The altered parameters, \tilde{m} and $\tilde{\sigma}$, can be evaluated by solving Eqs. 11 and 29 simultaneously. The application of the foregoing analyses will be shown in the following example, and Newton-Raphson's iteration scheme will be used to obtain the solution.

When the nonsymmetric fuzzy entropy function of Eq. 6 is used in Eq. 24, then the criterion for determining K in Eq. 26 becomes

$$\tilde{P}_i = \frac{1}{2}; f_i = 0 \dots \dots \dots (30)$$

$$\text{which leads to } K = \frac{\ln \frac{1}{2} + H}{\frac{1}{2} \ln \frac{1}{2}} \dots \dots \dots (31)$$

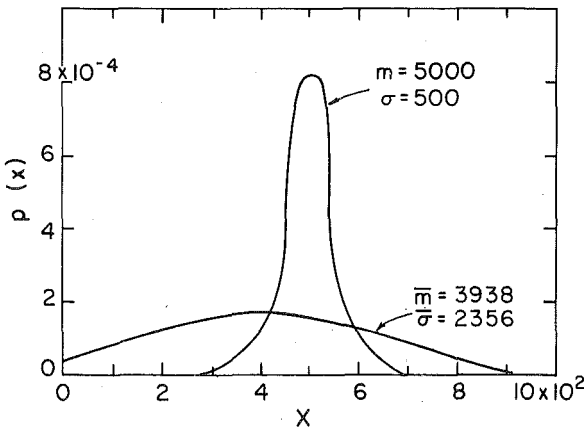


FIG. 2.—Alteration of Truncated Normal Distribution for Symmetrical Entropy (Eq. 4)—A of Fig. 1

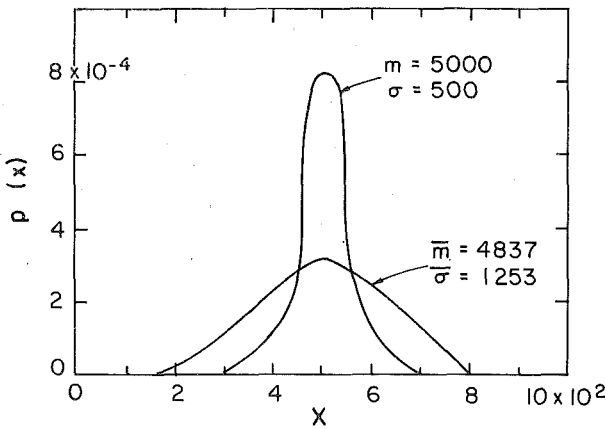


FIG. 3.—Alteration of Truncated Normal Distribution for Nonsymmetrical Entropy (Eq. 6)—B of Fig. 1

When the nonsymmetric fuzzy entropy function of Eq. 7 is used, then from the same criterion, Eq. 30, the multiplier K_L becomes

$$K_L = \frac{-\ln \frac{1}{2} - H}{2 \ln 2} \dots \dots \dots (32)$$

Example 2. The two parameters of a truncated normal variate, X , are $m = 5,000$ and $\sigma = 500$; and the range of X are $a = 0$ and $b = 10,000$. Let the fuzzy supports for the global uncertainty analysis be $f_G = 0.7$, and the fuzzy supports for the local uncertainty analysis be $f_1 = 0.5$ and $f_2 = 0.5$.

If the entropy function of Eq. 4 is used, then from Eqs. 11 and 29 the

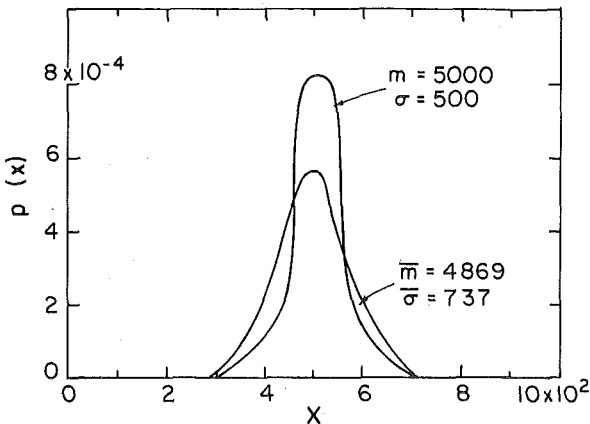


FIG. 4.—Alteration of Truncated Normal Distribution for Nonsymmetrical Entropy (Eq. 7)—C of Fig. 1

altered parameters are determined as $\bar{m} = 3,938$ and $\bar{\sigma} = 2,356$.

The altered distribution is shown in Fig. 2 with the original objective distribution. If the entropy function of Eq. 6 is used, the altered parameters become $\bar{m} = 4,837$ and $\bar{\sigma} = 1,253$ and the altered and original objective distributions are shown in Fig. 3.

If the entropy function of Eq. 7 is used, the altered parameters are $\bar{m} = 4,869$ and $\bar{\sigma} = 737$ and the altered and objective distributions are shown in Fig. 4.

REVIEW

Three elements of subjectivity have been presented in this paper:

1. The fuzzy support, f .
2. The fuzzy entropy function, G .
3. The criterion for determining the multiplier, K .

The fuzzy support, f , is a subjective evaluation of the factors which may alter the parameters σ and m . These factors can be decided on independently of the extent and the manner in which the parameters are changed. For instance if the parameters are obtained from cylinder test results and the parameters for in-situ structural concrete are altered from the cylinder test results by workmanship and weather, then, f is a measure of the subjective features workmanship and weather, but is not an evaluation of the effect of these features on σ and m . This effect is worked out by the aforementioned elements 2 and 3, the fuzzy entropy, G , and the criterion leading to K . This suggests that elements 2 and 3 have the characteristics of a fuzzy conditional relationship and that when the support, f , is numerically introduced fuzzy compositions result. A fuzzy relationship of the form; if the workmanship is α then the effect on the parameter is β would have linguistic statements for α and β . These statements would vary amongst different individuals. The curves in Fig. 1 also represent different fuzzy relationships between a fuzzy set and the parameter $\sigma/\bar{\sigma}$ which may be likened to the relationship between α and β . In this sense, the symmetrical entropy (A) has a range of $1/2 \leq f \leq 1$ with unbounded values of $\bar{\sigma}/\sigma$ at $f = 1/2$; whereas the nonsymmetrical cases (B and C) have the full range ($0 \leq f \leq 1$). For the same evaluation of the workmanship and weather (α), the effect on the parameters (β) is much smaller for the nonsymmetrical cases. The choice of the fuzzy entropy function, G , therefore establishes the investigators subjective opinion of the effect of workmanship and weather on the parameters.

The criteria by which K is established is a consequence of the entropy chosen. However, they essentially state that if the fuzzy support is at the lower limit then the investigator is unable to state that one proposition is more likely than another. This indifference statement is analogous to the principle of insufficient reason in probability theory.

An assessment procedure can now be suggested for the updating of probabilities by entropy procedures: (1) Determine the objective parameters m and σ ; (2) determine the fuzzy support for influencing features which can only be included subjectively (f_G); (3) use entropy functions A, B, or C depending on whether the influencing features are subjec-

tively assessed to have a big, medium or small effect on m or σ ; and (4) determine \bar{m} and $\bar{\sigma}$ Step 3 of this scheme is a general relationship between the subjective influencing features and the parameters in entropy form.

CONCLUSION

The uncertainty of a continuous variable is estimated by the combination of the objective entropy, which is obtained from statistics and the fuzzy entropy, which is a measure of imprecision. The global and local uncertainty analyses are introduced in determining the altered parameters of a probability distribution. Several fuzzy entropy functions are presented and their effects on the altered parameters are illustrated by two examples.

ACKNOWLEDGMENT

This study was supported by the National Science Foundation (Grant No. CME-8026583).

APPENDIX I.—REFERENCES

1. Brown, C. B., "Entropy Constructed Probabilities," *Journal of the Engineering Mechanics Division*, ASCE, Vol. 106, No. EM4, Proc. Paper 15635, Aug., 1980, pp. 633-640.
2. DeLuca, A., and Termini, S., "A Definition of a Nonprobabilistic Entropy in the Setting of Fuzzy Sets Theory," *Information and Control*, Vol. 20, 1972, pp. 301-312.
3. Munro, J., "Uncertainty and Fuzziness in Engineering Decision-Making," *Proceedings of the First Canadian Seminar on Systems Theory for the Civil Engineer*, Calgary, Alberta, Canada, Oct., 1979, pp. 18-19.
4. Shannon, C. E., "The Mathematical Theory of Communication," University of Illinois Press, Urbana, Illinois, 1949.
5. Zadeh, L. A., "Fuzzy Sets," *Information and Control*, Vol. 8, 1965.

APPENDIX II.—NOTATION

The following symbols are used in this paper:

- E = combined or total entropy;
- F = cumulative probability distribution;
- G = fuzzy entropy;
- H = Shannon's objective entropy;
- K = fuzzy multiplier;
- P = probability;
- X = random variable;
- f = fuzzy support;
- m = mean;
- p = probability density function; and
- σ = standard deviation.