The Third AFSS(1998), 696−700

Piecewise Linear Fuzzy Random Variables and their Statistical Application

Norio WATANABE[a], Tadashi IMAIZUMI[b]

[a] Department of Industrial and Systems Engineering, Chuo University
Kasuga, Bunkyo-ku, Tokyo 112, JAPAN
Fax: +81-3-3817-1943 E-mail: watanabe@indsys.chuo-u.ac.jp
[b] Tama University
Hijiriga-oka, Tama-shi, Tokyo 206, JAPAN

Abstract

Fuzzy random variables with piecewise linear membership functions are introduced from a practical viewpoint. The estimation of the expected values of these fuzzy random variables is also discussed and statistical application is demonstrated by using a real data set.

Keywords: fuzzy random variable, expectation, estimation

1. Introduction

The mathematical basis of fuzzy random variables was established by Puri and Ralescu [1] and other authors (cf. [3]). Though some authors have considered applications of fuzzy random variables (cf. [3]), applying fuzzy random variables in data analysis have not become common. One reason is that handling fuzzy random variables is not easy in practice. In this note we introduce piecewise linear fuzzy random variables which can be handled easily.

Piecewise linear fuzzy random variables in this note are formulated by piecewise linear membership functions which have turning points with the same membership values. The expected values of piecewise linear fuzzy random variables can be estimated by a nonparametric or parametric approach.

We can use piecewise linear fuzzy random variables for approximation of general fuzzy random variables and we will show an example. In this example we estimate the average meaning of a word.

Piecewise linear fuzzy random variables are applicable in data analysis and suitable for simulation studies because of facility in handling.

2. Piecewise linear fuzzy random variables

A fuzzy random variable is a mapping from the probability space (Ω, B, P) into the set Γ which consists of fuzzy sets. In this note we consider the fuzzy sets defined on the interval $I_X = [X_L, X_R]$, where X_L and X_R can take the values $-\infty$ and ∞ respectively.

Now we define the piecewise linear fuzzy random variables by considering the mapping from I_X into $[0,1]$. First divide the interval $[0,1]$ into $K-1$ sub-intervals $[\alpha_k, \alpha_{k-1}]$, where $0 = \alpha_K < \alpha_{K-1} < \cdots < \alpha_1 = 1$. Let x_0, $l_1, l_2, ..., l_K$, $r_1, r_2, ..., r_K$ be real valued random variables such that $l_k \geq 0$, $r_k \geq 0$. We obtain the piecewise linear line graph, which depends on a sample $\omega \in \Omega$, by connecting the points $(\min\{X_L, x_K\}, 0)$, (x_K, α_K), $(x_{K-1}, \alpha_{K-1}), ..., (x_1, \alpha_1)$, $(y_1, \alpha_1), ...,$ (y_K, α_K), $(\max\{X_R, y_K\}, 0)$ with straight lines sequentially, where $x_k = x_0 - \sum_{i=1}^{k} l_i$, $y_k = x_0 + \sum_{i=1}^{k} r_i$ for $k = 1, ..., K$ (see Figure 1).

Assuming that $l_k > 0$ if $x_{k-1} \in [X_L, X_R]$

and $r_k > 0$ if $y_{k-1} \in [X_L, X_R]$ for $1 \leq k \leq K$, we can define the membership function $\mu(x_0, l_1, ..., l_K, r_1, ..., r_K)$ by limiting the domain of the above line graph to $[X_L, X_R]$. Then we have the following result (cf. [3]).

Theorem 1. Let $X(\omega)$ be the fuzzy set defined by the membership function $\mu_{X(\omega)} = \mu(x_0(\omega), l_1(\omega), ..., l_K(\omega), r_1(\omega), ..., r_K(\omega))$ for $\omega \in \Omega$. Then, X, which is a mapping from Ω into Γ, is a fuzzy random variable.

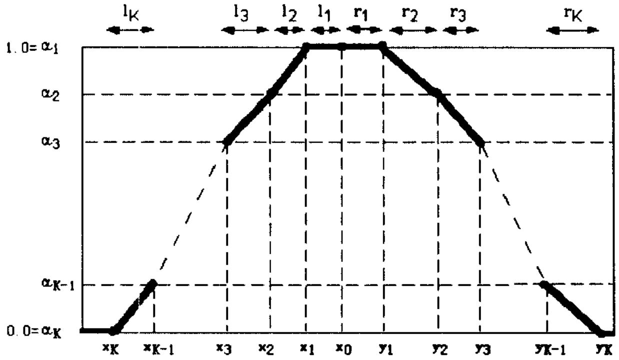

Figure 1. A piecewise linear membership function

We call X in Theorem 1 the piecewise linear fuzzy random variable. Clearly, Γ is a family of convex fuzzy sets which have continuous membership functions.

We can consider various types of membership functions of fuzzy random variables, by imposing some conditions on x_0, l_1,..., l_K, r_1,..., r_K. For example, linear, triangular or trapezoidal membership functions can be realized as shown in Figure 2. A trapezoidal membership function in Figure 2 (a) is obtained by setting $K = 2$. Moreover, by putting $l_1 = r_1 = 0$ (i.e. $x_1 = y_1 = x_0$), a triangular membership function (b) is obtained. Or, by putting $l_1 = 0$ and $r_1 = \infty$ we have a linear membership function (c).

Random variables x_0, l_1,..., l_K, r_1,..., r_K can be distributed in many ways. The following is an example of triangular fuzzy random variables.

Example 1. For $K = 2$ with $l_1 = r_1 = 0$, assume that x_0 is normally distributed and that l_2 and r_2 are distributed according to the lognormal distributions. Then X is a fuzzy random variable whose membership function is triangular.

The simplest example may be a fuzzy random variable satisfying the assumption that x_0 is normally distributed and l_2 and r_2 are positive constants in Figure 2 (b).

The following is another example.

Example 2. Provided that x_0 is a random variable or a constant, and that l_i and r_i $(i = 1, ..., K)$ are given by equations:

$$l_i = b_0^L + b_1^L(1 - (\alpha_{i-1} + \alpha_i)/2)^2,$$

$$r_i = b_0^R + b_1^R(1 - (\alpha_{i-1} + \alpha_i)/2)^2,$$

where b_0^L, b_0^R, b_1^L and b_1^R are positive random variables. Then X is a fuzzy random variable whose membership function is approximately quadratic.

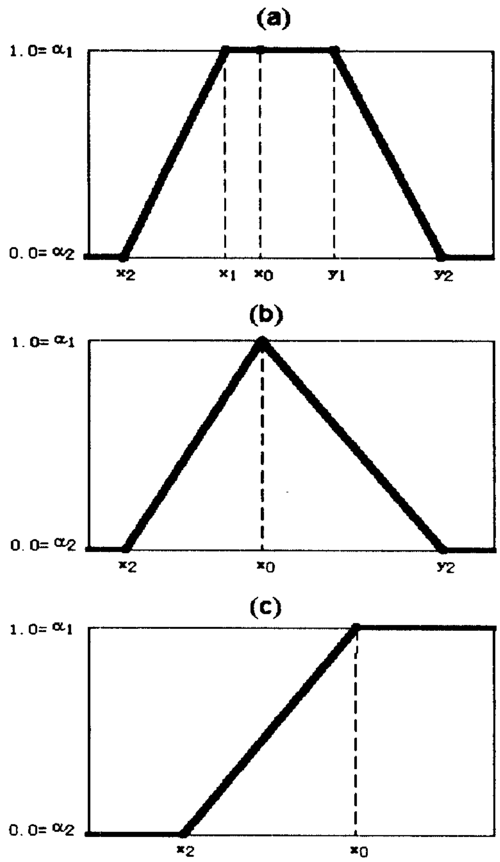

Figure 2. Examples of piecewise linear fuzzy random variables

For instance, set $K = 20$ and $x_0 = 0$, and assume that $b_0 \sim U[0.4, 0.6]$ and $b_1 \sim U[8, 10]$, where $U[\cdot, \cdot]$ stands for the uniform distribution. The bold curve in Figure 3 is one realization of the fuzzy random variable defined in Example 2.

The thin curve in Figure 3 is obtained by adding a multiplicative noise as follows:

$$l_i = (b_0^L + b_1^L(1 - (\alpha_{i-1} + \alpha_i)/2)^2)\exp[u_i],$$

$$r_i = (b_0^R + b_1^R(1 - (\alpha_{i-1} + \alpha_i)/2)^2)\exp[v_i],$$

where u_i and v_i are distributed according to $U[-0.75, 0.75]$.

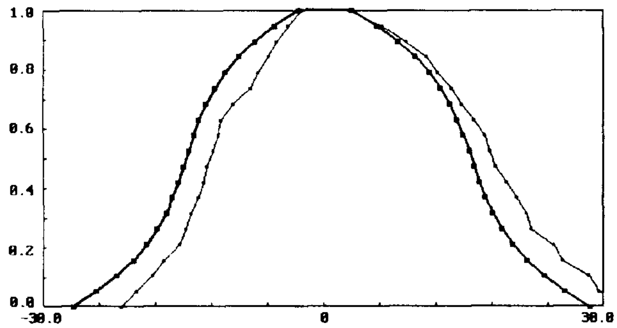

Figure 3. An example of a piecewise linear fuzzy random variable

3. Expectation

Let EX be the expected value of X and $(\cdot)_\alpha$ denote the $\alpha-$level set of a fuzzy set. We can prove the following theorem.

Theorem 2. Provided that Ex_0, El_k, Er_k $(k = 1, ..., K)$ exist. Then, for $\alpha_{k-1} \geq \alpha \geq \alpha_k$ $(2 \leq k \leq K)$,

$$(EX)_\alpha = [\ Ex_0 - \sum_{i=1}^{k-1} El_i - El_k \frac{\alpha_{k-1} - \alpha}{\alpha_{k-1} - \alpha_k},$$

$$Ex_0 + \sum_{i=1}^{k-1} Er_i + Er_k \frac{\alpha_{k-1} - \alpha}{\alpha_{k-1} - \alpha_k}\ ].$$

Note that $Ex_k = Ex_0 - \sum_{i=1}^{k} El_i$ and $Ey_k = Ex_0 + \sum_{i=1}^{k} Er_i$.

When a random sample $\{X_1, X_2, , ..., X_N\}$ is drawn from the population determined by X, the expected value EX $(= EX_n)$ can be estimated easily based on the expression in Theorem 2. If fuzzy random variables are parametric piecewise linear fuzzy random variables like examples as will be stated in the next section, then the parametric approach can be applied for estimation. In general cases the nonparametric approach is available. The nonparametric approach leads the estimate of EX satisfying the equation:

$$(\widehat{EX})_{\alpha_k} = [\sum_{n=1}^{N} x_{kn}/N, \sum_{n=1}^{N} y_{kn}/N],$$

where x_{kn} and y_{kn} are the lower and upper bounds of α_k-level set of X_n. An example is demonstrated in the section 5.

Facility of treating the expectation of piecewise linear fuzzy random variables is due to the assumption that a membership value of any turning point of membership functions is one of $\alpha_1, ..., \alpha_K$.

4. Parametric models

When the joint distribution of $\{x_0, l_1, l_2, ..., l_K, r_1, r_2, ..., r_K\}$ is determined by a few unknown parameters, we can say that the fuzzy random variable is parametric.

In Example 1, let express as $x_0 \sim N(m, \sigma^2)$, $\log l_2 \sim N(m_L, \sigma_L^2)$ and $\log r_2 \sim N(m_R, \sigma_R^2)$. Moreover we assume that x_0, l_2 and r_2 are independent each other. Then the parameter vector $\theta = (m, \sigma, m_L, \sigma_L, m_R, \sigma_L)$ determines the joint distribution.

In Example 2, the parameter vector consists of parameters of x_0, b_0 and b_1.

When a random sample $\{ X_1, ..., X_N \}$ is drawn, random samples $\{x_{01}, ...x_{0N}\}$, $\{l_{i1}, ...l_{iN}\}$, and $\{r_{i1}, ...r_{iN}\}$ are observed. Therefore unknown parameters can be estimated by a standard technique, for example, the maximum likelihood method. The expected value is estimated by substituting estimates into unknown parameters, if $(EX)_\alpha$ is represented by parameters explicitly.

We consider Example 1 with the parameter vector $\theta = (m, \sigma, m_L, \sigma_L, m_R, \sigma_L)$. From the relation between the mean of the lognormal distribution and the one of the corresponding normal distribution we can obtain the equation:

$$(EX)_\alpha = [m - (1-\alpha)\exp[m_L + \sigma_L^2/2],$$

$$m + (1-\alpha)\exp[m_R + \sigma_R^2/2]\ ]$$

for $0 \leq \alpha \leq 1$. We have an estimate of $(EX)_\alpha$ by substituting the sample means and sample covariances of x_0, $\log l_2$ and $\log r_2$ into θ. We note that a nonparametric approach is also applicable for estimating EX.

Parametric piecewise linear fuzzy random variables are suitable for simulation studies, where the randomly fluctuated fuzzy sets are required.

5. Statistical application

Watanabe [2] and Watanabe and Imaizumi [4] proposed estimation methods for the membership function. In their methods, membership functions of fuzzy random variables have limited forms. Use of piecewise linear fuzzy random variables relaxes this limitation, while calculation does not become so complicated.

In general, it is comparatively easier to observe membership values for given points in I_X than to observe X-values for given membership values α's. Given points and corresponding observed α-values determine a fuzzy set. A piecewise linear fuzzy set can approximate this fuzzy set, if K is sufficiently large. Thus piecewise linear fuzzy random variables can be used for approximation. The increase of K does not cause much difficulty in calculation.

Now we demonstrate an example of statistical application.

In this example the purpose is to know the average meaning of some word or some concept. The average meaning is represented by the expected value EX_n of individual meanings X_n. See [2] and [4] for details.

A questionnaire was conducted on a group of undergraduate students. The number of students is 100 (= N). Questions are about the word "tall" which is restricted to be used for the height of Japanese young male. Examinees were urged to evaluate membership values of the word "tall" for 11 points 160.0, 162.5, ..., 185 in centimeters. Evaluation was limited to select one α value among 0.0, 0.25, 0.5, 0.75, 1.0. Part of data is shown in Figure 4 by connecting the observed points, where each curve is shifted slightly. We achieved the approximation by piecewise linear fuzzy random variables with $K = 100$, $l_1 = 0$ and $r_1 = \infty$, and calculated the estimate of EX_n by the nonparametric approach. The estimated expected value is shown by the central line with dots in Figure 5.

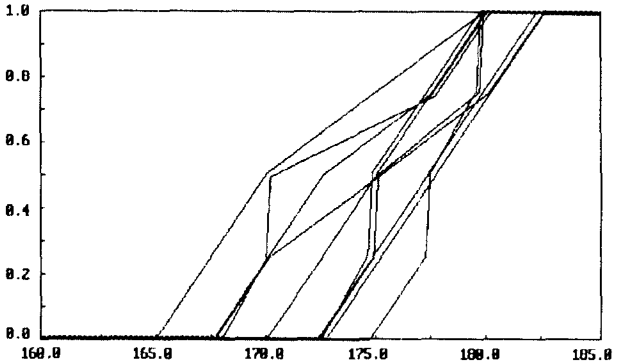

Figure 4. Part of a random sample

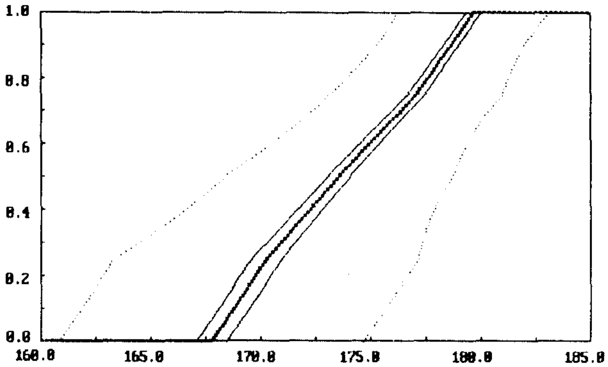

Figure 5. An estimate of the expected value

The interval between two thin lines for any fixed α in Figure 5 is the approximated 95% confidence interval of the lower bound of $(EX_n)_\alpha$. The confidence interval is based on the central limit theorem (see [3]). The interval between two dotted lines for a fixed α is the approximated 95%

tolerance interval of the lower bound of $(X_n)_\alpha$. The tolerance interval makes sense when the lower bound of $(X_n)_\alpha$ is normally distributed approximately. Since the number of α-values is limited, bends of thin lines and dotted curves is caused by the inequality:

$$SD(tU + (1-t)V) \leq tSD(U) + (1-t)SD(V)$$

for any random variables U and V and any real number $t \in [0, 1]$, where the equality holds when $t = 0$ and $t = 1$. SD stands for the standard deviation.

From Figure 5, we find the average meaning of the word "tall" in this group. Moreover we can see the magnitude of the fluctuation of the meaning.

The above example shows the applicability of piecewise linear fuzzy random variables.

References

[1] Puri, M. L. and Ralescu, D. A.: Fuzzy random variables, *J. of Math. Anal. and Appl.* 114 (1986), pp. 409-422.

[2] Watanabe, N.: Statistical method for estimating membership functions, *J. of Japan Soc. for Fuzzy Theory and Systems* **5** (1993), pp. 816-826 (in Japanese).

[3] Watanabe, N.: Fuzzy random variables and statistical inference, *J. of Japan Soc. for Fuzzy Theory and Systems* **8** (1996), pp.908-917.

[4] Watanabe, N. and Imaizumi, T.: Note on estimation of a membership function in a group, *Proceedings of First Asian Fuzzy Systems Symposium* (1993), pp. 636-641.