Proceedings of the 9th WSEAS International Conference on Applied Mathematics, Istanbul, Turkey, May 27-29, 2006 (pp309-314)

Fuzzy Boundaries of Sample Selection Model

L. MUHAMAD SAFIIH [1], ANTON ABDULBASAH KAMIL[2], M. T. ABU OSMAN [3]

[1,3]Mathematics Department

Faculty of Science and Technology, KUSTEM

21030 Kuala Terengganu, Terengganu, MALAYSIA.

http://www.kustem.edu.my

[2] School of Mathematical Sciences

Universiti Sains Malaysia

11800 USM Penang, MALAYSIA.

http://www.usm.my

Abstract-The issue of uncertainty and ambiguity are the major problem and become complicated in the modeling of econometrics. In this study, we will focus in the context of fuzzy boundaries of the sample selection model. The best approach of accounting for uncertainty and ambiguity is to take advantage of the tools provided by the theory of fuzzy sets. It seems particularly appropriate for modeling vague concepts. Fuzzy sets theory and its properties through the concept of fuzzy number and linguistic variables provide an ideal framework in order to solve the problem of uncertainty data. In this concept paper, we introduce a fuzzy membership function for solving uncertainty data of a sample selection model.

Key-Words:- uncertainty, ambiguity, sample selection model, crisp data, membership function

1 Introduction

The purpose this concept paper is to introduce a membership function of a sample selection model that can be used to deal with sample selection model problems in which historical data contains some uncertainty. With this concept provides an ideal framework to deal with problems in which there does not exist a definite criterion for discovering what elements belongs or do not belongs to a given set (Miceli, 1998)

Fuzzy set defines by a fuzzy sets in a universe of discourse U is characterized by a membership function denoted by the function μ_A maps all elements of U that take the values in the interval [0,1] that is $A: X \longrightarrow [0,1]$ (Zadeh, 1965)

Clearly, a fuzzy set is a generalization of the concept of a set whose membership function takes only two values: $\{0,1\}$ and in principle any function of the form $A: X \longrightarrow [0,1]$ describe a membership function associated with a fuzzy sets.

In this paper, we use a monotonically increasing *S-Sharp* (*S-Curve*) membership function as represented by the generic graphical representation in Fig. 1.

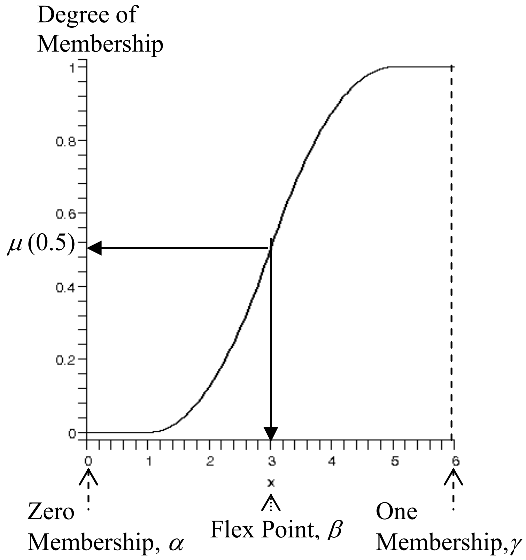

Fig.1 The *S*-Curve Function Characteristics

From Figure 1, we can see clearly the characteristic of the *S*-Curve. It's zero membership value (α), it's complete membership value (γ) and a third piece of information, the parameter β, $\beta = \frac{\alpha + \gamma}{2}$, is the inflection or crossover point. The inflection point is chosen by the analyst according to the apparent, suspected, or known distribution of the population (Cox, 1998). This is the point at which the domain is 50% true.).

An *S-Curve* is defined using three parameters, α, β and γ (*Cox*, 1998). The membership function as proposed by Zadeh (1965) can be express as:

$$F(x,\alpha,\beta,\gamma)=\begin{cases} 0 & x\leq\alpha \\ 2((x-\alpha)/(\gamma-\alpha))^2 & \alpha\leq x\leq\beta \\ 1-2((x-\gamma)/(\gamma-\alpha))^2 & \beta\leq x\leq\gamma \\ 1 & x\geq\gamma \end{cases} \quad (1)$$

2 Fuzzy Sets and Fuzzy Numbers

According Fuzzy logic is the superset of classical logic with the introduction of "degree of membership". The introduction of degree of membership allows the input to interpolate between the crisp set (*Jamshidi*, *et.al*., 1993). Unlike classical set theory that classifies the elements of the set into crisp sets; fuzzy set has an ability to classify elements into a continuous set using the concept of degree of membership. The characteristic function or membership function not only gives 0 or 1 but can also give values between 0 and 1. In other word, to express a number in words, we need a way to translate input numbers into confidences in a fuzzy set of word descriptors, the process of fuzzification (or the process of converting numerical measurements to confidences (grades of membership) of fuzzy set members is called fuzzification). In fuzzy math, that is done by membership functions

In this paper we consider the sample selection model proposed by Heckman (1979). The conventional sample selection model has the form

$$\begin{aligned} &Y_i^* = \gamma_0 + Z_i'\vartheta_0 + u_i \\ &D_i = 1 \quad if \ D_i^* = X_i'\beta + \varepsilon_i > 0 \\ &D_i = 0 \ otherwise \quad i = 1,\dots,n \\ &Y_i = Y_i^* D_i \end{aligned} \quad (2)$$

where, D_i, Y_i are the dependent variables, X and Z are vectors of exogenous variables, γ_0, β and ϑ are unknown parameter vectors, and u_i and ε_i are zero mean error terms. The first equation is the outcome equation and the second equation is the participation equation. Without loss of generality, assume all observations are fuzzy numbers, since crisp values can be represented by degenerated fuzzy number. Then, the fuzzy sample selection model is given by:

$$\begin{aligned} &\tilde{Y}_i^* = \gamma_0 + Z_i'\tilde{\vartheta}_0 + \tilde{u}_i \\ &D_i = 1 \quad if \ D_i^* = X_i'\tilde{\beta} + \tilde{\varepsilon}_i > 0 \\ &D_i = 0 \ otherwise \quad i = 1,\dots,n \\ &Y_i = Y_i^* D_i \end{aligned} \quad (3)$$

where $\tilde{Y}_i^*, \tilde{\vartheta}_o, \tilde{\beta}, \tilde{u}_i$ *and* $\tilde{\varepsilon}_i$ are fuzzy numbers with the membership functions, $\mu_{\tilde{Y}_i^*}, \mu_{\tilde{\vartheta}_o}, \varpi_{\tilde{\beta}}, \mu_{\tilde{u}_i}$ *and* $\mu_{\tilde{\varepsilon}_i}$, respectively. The problem is to find the estimates for γ_0 $\tilde{\vartheta}_0$, $\tilde{\beta}$, $\tilde{u}_i$ and $\tilde{\varepsilon}_i$ consistently from the data { Y_i, D_i, X_i and Z_i), i =1,..*n*.

2.1 Arithmetic operations on FN

The Zadeh's extension principle method and the α-cut method are two methods that can be applied to define arithmetical operations with fuzzy numbers. Both methods are very important concept in the relationship between fuzzy sets and crisp sets. Zadeh extension principle provides a general method for extending the classical (traditional) and crisp mathematical concept and arithmetic operations to the fuzzy environment while in the α-cut method, the arithmetic for fuzzy number can be reduced to interval arithmetic.

2.1.1 Linguistic Variable

Linguistics variables are concept which plays an important role in applications of fuzzy sets. Informally, a linguistic variable, ζ is a variable whose values are words or sentences which serve as names of fuzzy subsets of a universe of discourse or variable whose values are not numbers but words or sentences in a natural or artificial language (Zimmermann, 1985). According to Zadeh (1974), in general, verbal characterizations are less precise then numerical ones, and thus serve the function of providing a means of approximate description of phenomena which are too complex or too ill-defined to admit of analysis in conventional quantitative terms. The motivation for this concept derives from the observation that in most of our commonsense reasoning, we employ words rather than numbers to described the values of variables (Zadeh, 1987)

3 Fuzzy Boundaries of Sample Selection Model

Definition 3.1 Let U be the universe of discourse, $U = \{u_1, u_2, \cdots, u_n\}$. A fuzzy set A of U is defined by

$$A = \frac{f_A(u_1)}{u_1} + \frac{f_A(u_2)}{u_2} + \cdots + \frac{f_A(u_n)}{u_n} \qquad (4)$$

where f_A is the membership function of A, $f_A : U \to [0,1]$, and $f_A(u_i)$ indicates the grade of membership of u_i in A, where $f_A(u_i) \in [0,1]$ and $1 \le i \le n$.

Instead of using crisp value i.e., single-valued parameters on which statistical inference may be done in the case of classical sample selection model, fuzzy parameter in the form of membership function are used. From the above definition we know that the membership function is constructed from the quadratic function. In order to construct the fuzzy boundaries (minimum and maximum), we consider the following function:

$$\tilde{F}(x,\alpha,\beta,\gamma) = \begin{cases} 0, & 0 \le x \le \alpha \\ f := a_1x^2 + b_1x + c_1, & \alpha \le x \le \beta \\ g := a_2x^2 + b_2x + c_2, & \beta \le x \le \gamma \\ 1, & x \ge \gamma \end{cases} \qquad (5)$$

Then we need to determine the value of a_i, b_i and c_i where $i = 1,2$. And it's should meet and satisfy the conditions (1) $g_1(\alpha) = 0$, (2) $g_1(\beta) = 1/2$, (3) $g_1{}'(\alpha) = 0$, (4) $g_2(\gamma) = 1$, (5) $g_2(\beta) = 1/2$, and (6) $g_2{}'(\gamma) = 0$. The function $\tilde{F}$ should continue and following S curve term. Here, we propose that the functions g_1 and g_2 meet at the point $(\beta, 1/2)$. Define $\beta = \frac{w_1\alpha + w_2\gamma}{w_1 + w_2}$, where w_1 and w_2 are the weighted. The weighted values are chosen, it is depend on the values of maximum and minimum boundaries we need. When, $w_1 = w2$, we get the middle membership function, as in (4). When, $w_1 > w2$, we get the upper (maximum) membership function, $\mu'(x)$, and when, $w_1 < w2$, we get the lower (minimum) membership function, $\mu(x)$. We consider $w_1 = 2$ and $w_2 = 4$ to construct the minimum membership function $\mu(x)$**.** This follows that $\beta = \frac{2\alpha + 4\gamma}{6}$. We then substitute

$\beta = \dfrac{2\alpha + 4\gamma}{6}$ into the function $\widetilde{F}$ which satisfies the six conditions and solve them to get a_i, b_i and c_i. After simplifying we get the minimum membership function $\mu(x)$ as follows

$$\widetilde{F}_{\min}(u,\alpha,\beta,\gamma) = \begin{cases} 0, & 0 \le u \le \alpha \\ \dfrac{9}{8}\left(\dfrac{u-\alpha}{\alpha-\gamma}\right)^2, & \alpha \le u \le \beta \\ -\dfrac{9}{2}\left(\dfrac{u-\gamma}{\alpha-\gamma}\right)^2 + 1, & \beta \le u \le \gamma \\ 1, & u \ge \gamma \end{cases} \quad (6)$$

We consider $w_1 = 4$ and $w_2 = 2$ to construct the maximum membership function $\mu'(x)$. Here, we have $\beta = \dfrac{4\alpha + 2\gamma}{6}$. After doing similar procedure as above, we get maximum membership function $\mu'(x)$ as follows:

$$\widetilde{F}_{\max}(u,\alpha,\beta,\gamma) = \begin{cases} 0, & 0 \le u \le \alpha \\ \dfrac{9}{2}\left(\dfrac{u-\alpha}{\alpha-\gamma}\right)^2, & \alpha \le u \le \beta \\ -\dfrac{9}{8}\left(\dfrac{u-\gamma}{\alpha-\gamma}\right)^2 + 1, & \beta \le u \le \gamma \\ 1, & u \ge \gamma \end{cases} \quad (7)$$

4 Result

The maximum and minimum membership functions depend on the chosen value of w_1 and w_2. When the value of the weighted is quite similar, then the membership function is close to the middle membership function. In the following graph of (4), (5), and (6), we take $\alpha = 1$, $\gamma = 5$.

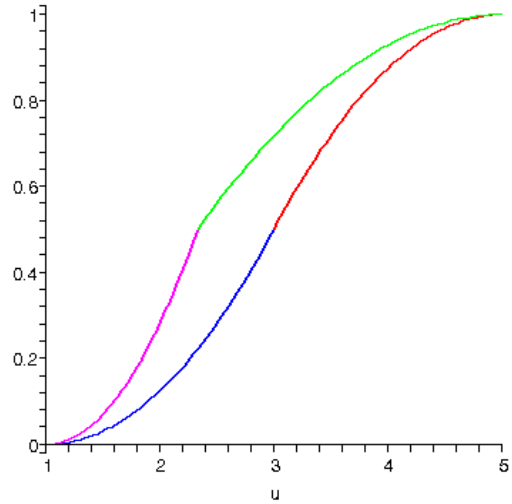

Fig.2 The maximum boundary of *S*-Curve membership function

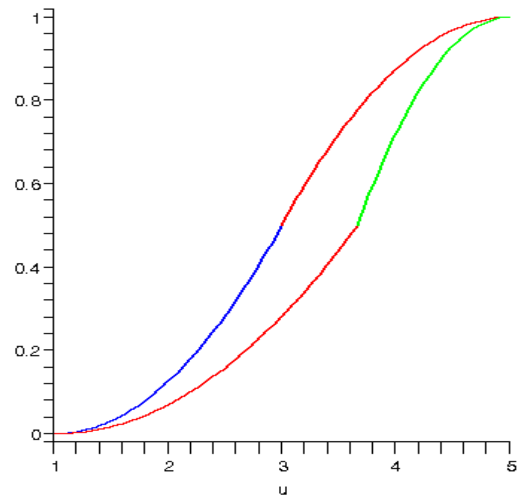

Fig.3 The minimum boundary of *S*-Curve membership function

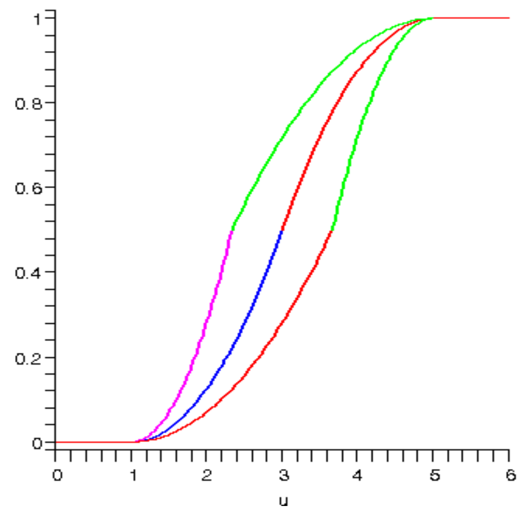

Fig.4 The crisp, minimum and maximum boundaries of *S*-Curve membership function

4 Conclusion

In this research, we introduced a fuzzy membership function of a sample selection model. In a sample selection model, the major problems of uncertainty data can be solved by using set of fuzzy number as a fuzzy boundaries. Instead of using crisp value i.e. single-valued parameters on which statistical inference in the case of classical sample selection model, the quadratic term in term of membership function are used. A fuzzy boundary of minimum and maximum for A Fuzzy Sample Selection Model *S*-Curve (FBSSM *S*-Curve) is define and clearly represented by using graphical.

References:

[1] Cox, E., *The Fuzzy System Handbook*, 2nd ed. AP Profesional, New York, 1998.

[2] Heckman, J.J., Sample selection as a specification error, *Econometrica*, Vol.47, 1979 pp.153-161.

[3] Jamshidi, M., Vadiee, N., Ross, T. J., *Fuzzy Logic and Control*, Prentice Hall, 1993

[4] Miceli, D., Measuring Poverty Using Fuzzy Sets, Discussion Paper no. 38, 1998, University of Canberra.

[5] Zimmermann, H.J.. *Fuzzy Sets Theory and its Applications*, M.A : Kluwer Academic, Boston, 1985.

[6] Zadeh, L.A., Fuzzy Sets, *Information and Control*, Vol.8(3), 1965, pp.338-353.

[7] Zadeh, L.A., On the analysis of large-scale system. In: Gottinger, H., ed., *Systems approaches and Environment Problems*. Vandenhoeck and Ruprecht, Gottingen, 1974.

[8] Zadeh, L.A., Fuzzy Sets: Usuality and Commonsense Reasoning, In: Vaina, L., ed., *Matters of Intelligence*, D. Reidel Publishing Company, Dordrecht, 1987.