ELSEVIER
Fuzzy Sets and Systems 94 (1998) 71–75
FUZZY sets and systems

Short Communication

# Note on maxmin $\mu/E$ estimation

Wang Xizhao, Ha Minghu*

*Department of Mathematics, Hebei University, Baoding, Hebei 071002, People's Republic of China*

Received January 1995; revised July 1996

**Abstract**

We present a solution to the Maxmin $\mu/E$ estimation problem of the family of fuzzy numbers with two parameters, location and scale, and show certain important properties of the Maxmin $\mu/E$ estimator. © 1998 Elsevier Science B.V.

*Keywords:* Possibility variables; Fuzzy numbers; Location parameter; Scale parameter

## 1. Introduction

Maxmin $\mu/E$ estimation for the family of fuzzy numbers with location and scale is a kind of method of estimating parameters for determining membership functions. Variations of this method has appeared in many papers (e.g. [4–8]). In paper [5], the estimation of the location parameter is first considered and the solution is given. In papers [6–8], the special case of this estimation is discussed by using the possibility theory. In paper [4], some theoretical results on this method are obtained by drawing possibilistic inferences. In this paper, we continue the discussion of the Maxmin $\mu/E$ estimation and present a solution to the family of fuzzy numbers with two parameters, location parameter and scale parameter, and show certain important properties. Throughout this paper, $\vee$ and $\wedge$ denote max and min, respectively.

First, in this section, we will briefly review the possibility theory and set up notations needed in the paper. Any fuzzy number $\mu$, with membership function $\mu(x)$, will be defined by (1) $\mu_\alpha = \{x: \mu(x) \geqslant \alpha\}$ is a bounded, closed interval for each $\alpha \in (0, 1)$, and (2) $\mu_1 = \{x: \mu(x) = 1\}$ is not empty. The set of all fuzzy numbers is denoted by $F(R)$. Let

$$L(\mu) = \left\{ Q_\mu(a, b) \mid Q_\mu(a, b)(x) = \mu\left(\frac{x-a}{b}\right),\ a \in R, b > 0 \right\},$$

where $\mu$ is a 0-symmetric fuzzy number (i.e. $\mu(-x) = \mu(x)$ for every $x \in R$), the parameter $a$ is called location and the parameter $b$ is called scale. By Zadeh's extension principle, we have

$$\sum_{i=1}^{n} c_i Q_\mu(a_i, b_i) = Q_\mu\left(\sum_{i=1}^{n} c_i a_i,\ \sum_{i=1}^{n} |c_i| b_i\right) \in L(\mu),$$

*Corresponding author.


*PII* S0165-0114(96)00245-X

where $c_i$ $(i = 1, 2, \ldots, n)$ are $n$ real numbers and $Q_\mu(a_i, b_i)$ $(i = 1, 2, \ldots, n)$ are in $L(\mu)$.

Let us now consider the possibility variables. Suppose $(U, P(U), \Pi)$ denotes a possibility measure space where $U$ is a set denoting all of psychological observations, $P(U)$ denotes the set of all subsets in $U$ and $\Pi$ is a possibility measure, and $R$ is the set of all real numbers. Any possibility variable $\xi$ is a map $U \to R$. Let $\mu_\xi(x) = \Pi(\{u: \xi(u) = x\})$ for $x \in R$; then $\mu_\xi$ represents the distribution function of the possibility variable $\xi$. It is clear that $\mu_\xi$ is a map from $R$ to $[0, 1]$. If $\mu_\xi$ is a fuzzy number then we say that $\xi$ is regular. We denote by

$$E[\xi] = \int_R \mu_\xi(t)\,\mathrm{d}t.$$

Then the value of $E[\xi]$ represents a measure of the fuzziness of the distribution $\mu_\xi$. It is easy to prove that $E[\xi] = lb$ if $\mu$ is a 0-symmetric fuzzy number, the distribution function of $\xi$ is $Q_\xi(a, b)$ and $\int_R \mu(t)\,\mathrm{d}t = l$.

We now define $\mu_{\xi_1\xi_2\cdots\xi_n}(x_1, x_2, \ldots, x_n)$, the joint distribution of a possibility vector $(\xi_1, \xi_2, \ldots, \xi_n)$ by $\Pi(\{u \mid \xi_1(u) = x_1, \xi_2(u) = x_2, \ldots, \xi_n(u) = x_n\})$. Possibility variables $\xi_1, \xi_2, \ldots, \xi_n$ are called independent if

$$\Pi(\{u \mid \xi_{n_1}(u) = x_{n_1}, \ldots, \xi_{n_k}(u) = x_{n_k}\}) = \bigwedge_{i=1}^{k} \Pi(\{u \mid \xi_{n_i}(u) = x_{n_i}\})$$

for all $(x_{n_1}, \ldots, x_{n_k}) \in R^k$ and $(n_1, \ldots, n_k) \subset \{1, 2, \ldots, n\}$. Obviously, if $f$ is a function defined on $R^n$ and $f(\xi_1, \xi_2, \ldots, \xi_n)(u)$ is defined by $f(\xi_1(u), \xi_2(u), \ldots, \xi_n(u))$ for all $u \in U$, then $f(\xi_1, \xi_2, \ldots, \xi_n)$ is also a possibility variable.

Our discussion in the following sections will employ certain possibilistic concepts. For a given type, all these classes of distribution functions depend on one or several parameters. The parameters may vary over a specified range, called parameter space. The collection of all the distribution functions, when the parameters vary over their possible range is called a family. Generally, the actual distribution is hidden in the family. Let $\xi_1, \xi_2, \ldots, \xi_n$ be $n$ independent, identical distribution possibility variables (i.i.d.); then $(\xi_1, \xi_2, \ldots, \xi_n)$ is called a sample of the family. The observed values of the sample may be obtained by psychological observation. Any function of the sample $(\xi_1, \xi_2, \ldots, \xi_n)$, which does not involve unknown parameters, is called a statistic. Obviously, every statistic is a possibility variable.

Let the family be such a class of distribution $\{\mu_\xi(x, \theta) \mid \theta = (\theta_1, \ldots, \theta_n) \in \Theta\}$ where $\Theta$ is parameter space ($\Theta \subset R^n$). We consider how to estimate reasonably the true value of the parameter $\theta$ after observing psychologically the family $m$ times and obtaining a observed sample $(x_1, x_2, \ldots, x_m)$. Let $E[\xi] = \int_R \mu_\xi(x, \theta)\,\mathrm{d}x$, $(\xi_1, \xi_2, \ldots, \xi_n)$, be a sample of the family and $(x_1, x_2, \ldots, x_m)$ an observed value of the sample. The joint distribution of the sample, which is regarded as the possibility with which the observed sample $(x_1, x_2, \ldots, x_m)$ appears, can be denoted by

$$\mu(x_1, \ldots, x_m, \theta) = \bigwedge_{j=1}^{m} \mu_\xi(x_j, \theta),$$

where $\theta = (\theta_1, \theta_2, \ldots, \theta_n)$. Noting that changes of values of parameters $\theta_1, \theta_2, \ldots, \theta_n$ directly influence the possibility that the observed sample $(x_1, x_2, \ldots, x_m)$ appears and the value of $E[\xi]$, we should select $\theta$ such that the possibility is as high as possible and the value of $E[\xi]$ is as small as possible. Therefore, we denote by

$$L(\theta) = \bigwedge_{j=1}^{m} \mu_\xi(x_j, \theta)/E[\xi]$$

for all $\theta = (\theta_1, \theta_2, \ldots, \theta_n) \in \Theta$. If there exists $\hat{\theta} \in \Theta$ such that $L(\hat{\theta}) = \max_{\theta \in \Theta} L(\theta)$ then $\hat{\theta}$ is called Maxmin $\mu/E$ estimator of the parameter $\theta$.

## 2. A solution

**Theorem 2.1.** *Let $\mu$ be a 0-symmetric fuzzy number, $L(\mu) = \{Q_\mu(a, b) \mid a \in R, b > 0\}$ be the family and $(x_1, x_2, \ldots, x_m)$ be an observed sample of the family. Then the Maxmin estimator of the parameter $\theta = (a, b)$ is*

$$(\hat{a}, \hat{b}) = ((x^{(1)} + x^{(m)})/2, (x^{(m)} - x^{(1)})/2c),$$

*where $x^{(1)} \leqslant x^{(2)} \leqslant \cdots \leqslant x^{(m)}$ are the ordered values of $x_1, x_2, \ldots, x_m$ $(m \geqslant 2)$, $c$ is a real number at which the function $g(t) = t\mu(t)$ $(t \geqslant 0)$ attains its maximum.*

**Proof.** Let $\int_{-\infty}^{\infty} \mu(t)\,dt = l$. It is easy to compute that $E[L_\mu] = lb$. Therefore,

$$\begin{aligned} L(a,b) &= \bigwedge_{j=1}^{m} \mu\left(\frac{x_j - a}{b}\right) \Big/ (lb) \\ &= \bigwedge_{j=1}^{m} \mu\left(\frac{x^{(j)} - a}{b}\right) \Big/ (lb) \\ &= \left[\mu\left(\frac{x^{(1)} - a}{b}\right) \wedge \mu\left(\frac{x^{(m)} - a}{b}\right)\right] \Big/ (lb). \end{aligned}$$

By dividing $R$ into three subintervals, $(-\infty, x^{(1)})$, $[x^{(1)}, x^{[m]}]$ and $(x^{(m)}, \infty)$, we obtain that $\mathrm{Max}_{a\in R} L(a,b)$ attains its maximum at $a = (x^{(1)} + x^{(m)})/2$ for any given $b > 0$; it can also be found in [1, Lemma 1]. Hence,

$$\begin{aligned} &L\left(\frac{x^{(1)} + x^{(m)}}{2}, b\right) \\ &= \frac{1}{lb} \bigvee_{a\in R} \left[\mu\left(\frac{x^{(1)} - a}{b}\right) \wedge \mu\left(\frac{x^{(m)} - a}{b}\right)\right] \\ &= \frac{1}{lb}\mu\left(\frac{x^{(m)} - x^{(1)}}{2b}\right). \end{aligned}$$

Let $t = (x^{(m)} - x^{(1)}/2b$. Then

$$L\left(\frac{x^{(1)} + x^{(m)}}{2}, b\right) = \frac{2t\mu(t)}{l(x^{(m)} - x^{(1)})}.$$

As the assumption that $t\mu(t)$ attains the maximum at $t = c$, $L((x^{(1)} + x^{(m)})/2, b)$ attains its maximum at $b = (x^{(m)} - x^{(1)})/2c$. Therefore, the Maxmin $\mu/E$ estimation of parameter $(a, b)$ is

$$(\hat{a}, \hat{b}) = ((x^{(1)} + x^{(m)})/2, (x^{(m)} - x^{(1)})/2c).$$

The proof is completed. □

**Example 2.1.** Let us now consider the Maxmin $\mu/E$ estimator of the scale parameter, $b$, of the family $Q_\mu(a, b)$. $(x_1, x_2, \ldots, x_m)$ is an observed sample of the family and $(x^{(1)}, x^{(2)}, \ldots, x^{(m)})$ is the ordered arrangement of the sample.

(a) If

$$\mu(t) = \begin{cases} 1 - |x|^k & \text{if } |x| < 1 \\ 0 & \text{otherwise} \end{cases} \quad (k > 0),$$

then $\hat{b} = \frac{1}{2}(k+1)^{1/k}(x^{(m)} - x^{(1)})$.

(b) If $\mu(t) = \exp(-\pi t^2)$, then $\hat{b} = \sqrt{\frac{1}{2}\pi}\,(x^{(m)} - x^{(1)})$.

(c) If $\mu(t) = \mathrm{e}^{-|t|}$, then $\hat{b} = \frac{1}{2}(x^{(m)} - x^{(1)})$.

(d) If

$$\mu(t) = \begin{cases} \sqrt{1 - t^2}, & |t| < 1, \\ 0, & |t| \geqslant 1, \end{cases}$$

then $\hat{b} = (1/\sqrt{2})(x^{(m)} - x^{(1)})$.

The observed sample in Theorem 2.1 is regarded as crisp observations, but fuzzy observations can also be used.

**Example 2.2.** Let $\mu(t) = 1 - |t|$ ($|t| < 1$) and $L(\mu) = \{\mu((x - a)/b) \mid a \in R, b > 0\}$. We consider the problem of estimating vague location and vague scale for given fuzzy data: $X_1 = 1 - |t - 20|/8$, $12 \leqslant t \leqslant 28$, $X_2 = 1 - |t - 30|/20$, $10 \leqslant t \leqslant 50$ and $X_3 = 1 - |t - 40|/8$, $32 \leqslant t \leqslant 48$. According to Zadeh's extension principle, we can evaluate $\mathrm{Max}(X_1, X_2, X_3)$ and $\mathrm{Min}(X_1, X_2, X_3)$. This results in

$$\tilde{\hat{a}} = \begin{cases} 1 - (t - 21)/14, & 21 \leqslant t \leqslant 70/3, \\ 1 - |t - 30|/8, & 70/3 \leqslant t \leqslant 110/3, \\ 1 - (39 - t)/14, & 110/3 \leqslant t \leqslant 39, \\ 0 & \text{otherwise} \end{cases}$$

$$\tilde{\hat{b}} = \begin{cases} 1 - (5t - 20)/64, & 4 \leqslant t \leqslant 20/3, \\ 1 - |t - 10|/16, & 20/3 \leqslant t \leqslant 100/3, \\ 1 - (190 - 5t)/16, & 100/3 \leqslant t \leqslant 38, \\ 0 & \text{otherwise} \end{cases}$$

by Theorem 2.1.

## 3. Properties

**Definition 3.1.** Let $F$ be a family of distribution functions, $\xi = (\xi_1, \xi_2, \ldots, \xi_n)$ be a sample, and let $T(\xi)$ be a statistic whose distribution function belongs to $F$. We say $T(\xi)$ is sufficient with respect to $F$ if the joint distribution function of $(T(\xi), \xi)$, $G(t, x)$, does not depend on $x$.

**Note 3.1.** Every sample includes a certain amount of information on the family. Definition 3.1 shows that a sufficient statistic contains same amount of information as the sample with respect to the family. It follows that a sufficient statistic can be used to simplify a sample without losing information.

**Definition 3.2.** Let $F$ be a family denoted by $F = \{F(x,\theta) \mid \theta \in \Theta\}$ where $\Theta$ is parameter space, $\bar{\xi} = (\xi_1, \xi_2, \dots, \xi_n)$ a sample of the family and $T(\bar{\xi})$ a statistic. We say

(a) $T(\bar{\xi})$ is a sufficient estimator of $\theta$ if $T(\bar{\xi})$ is sufficient with respect to $F$;

(b) $T(\bar{\xi})$ is a consistent estimator of $\theta$ if $\Pi(\{T(\bar{\xi}) = \theta\}) = 1$;

(c) $T(\bar{\xi})$ is a maximum likelihood estimator of $\theta$ if

$$L(x_1, \dots, x_n, T(\bar{\xi})) = \operatorname*{Max}_{\theta \in \Theta} L(x_1, \dots, x_n, \theta)$$

where $L(x_1, \dots, x_n, \theta) = \bigwedge_{j=1}^{n} F(x_j, \theta)$.

**Note 3.2.** Sufficiency shows that the estimator can be used to estimate the parameter without losing information. Consistency illustrates the possibility that the estimator takes true value is maximum. Maximum likelihood explains the possibility that the sample appears attains maximum. Therefore, Definition 3.2 may be regarded as a criterion for judging reasonableness of an estimator.

**Theorem 3.1.** *Let the family be $L(\mu)$, $(\xi_1, \xi_2, \dots, \xi_n)$ be a sample of the family, $(x_1, x_2, \dots, x_n)$ an observed value of the sample, $m = \min(\xi_1, \xi_2, \dots, \xi_n)$, $M = \operatorname{Max}(\xi_1, \xi_2, \dots, \xi_n)$, $\bar{X} = (M + m)/2$ and let $S = (M - m)/2c$ where $c$ is a number at which the function $t$, $\mu(t)$, attains maximum. Then*

*(a) $\bar{X}$ is a consistent estimator of the parameter $a$ and is a maximum likelihood estimator of the parameter $a$ when $b$ is known;*

*(b) $(\bar{X}, S)$ is a sufficient estimator of the parameter vector $(a, b)$.*

In order to prove the theorem, we need the following lemmas.

**Lemma 3.1.** *Any symmetric fuzzy number $Q_\mu(x; a, b)$ has a membership function as*

$$Q_\mu(x; a, b) = \begin{cases} \mu(-(x-a)/b) & \text{if } x < a, \\ 1 & \text{if } x = a, \\ \mu((x-a)/b) & \text{if } x > a, \end{cases}$$

*where $\mu(x)$ is monotonically decreasing on $(0, \infty)$, left continuous, $0 \leqslant \mu(x) \leqslant 1$, and $\lim_{x \to \infty} \mu(x) = 0$.*

**Proof.** See [2]. □

**Lemma 3.2.**

$$\bigvee_{\Gamma} \left( Q_\mu(x; a, b) \wedge \bigwedge_{j=2}^{n} Q_\mu(y_j; a, b) \right) = Q_\mu(x; a, b)$$

*and*

$$\bigvee_{\Gamma_1} \left( Q_\mu(x; a, b) \wedge Q_\mu(y; a, b) \wedge \bigwedge_{j=2}^{n-1} Q_\mu(t_j; a, b) \right)$$
$$= Q_\mu(x; a, b) \wedge Q_\mu(y; a, b)$$

*where*

$$\Gamma = \{(y_2, \dots, y_n) \in R^{n-1} \mid y_j \leqslant x, j = 2, \dots, n\}$$

*and*

$$\Gamma_1 = \{(t_2, \dots, t_{n-1}) \in R^{n-2} \mid x \leqslant t_j \leqslant y, \; j = 2, \dots, n-1\}.$$

**Proof.** The first equality can be obtained from the fact that the inequality

$$\bigvee_{\Gamma} (Q_\mu(x; a, b) \wedge Q_\mu(y_j; a, b)) \leqslant Q_\mu(x; a, b)$$

holds and the sign of equality holds if $y_2 = y_3 = \cdots = y_n = x$; the other equality is a direct result of the monotonicity of $Q_\mu$. □

**Lemma 3.3.** Let *$\xi_1, \xi_2, \dots, \xi_n$ be $n$ i.i.d. possibility variables having the same distribution function $Q_\mu(x; a, b)$, $M = \max(\xi_1, \xi_2, \dots, \xi_n)$ and let $m = \min(\xi_1, \xi_2, \dots, \xi_n)$. Then $M$ and $m$ are also two i.i.d. possibility variables having the same distribution function as $\xi_1$.*

**Proof.** The assertion that $M$ and $m$ have the same distribution functions as $\xi_1$ is a consequence of

Theorem 1 in [4]. It remains to show the independence. By Lemma 3.2 we have

$$\Pi(\{m = x, M = y\})$$
$$= \Pi\left(\bigcup_{\Gamma_1} \{\xi_{i_1} = x, \xi_{i_n} = y, \xi_{i_j} = t_j, j = 2, \dots, n-1\}\right)$$
$$= \bigvee_{\Gamma_1} \Pi(\{\xi_{i_1} = x, \xi_{i_n} = y, \xi_{i_j} = t_j, j = 2, \dots, n-1\})$$
$$= \bigvee_{\Gamma_1}\left(\Pi(\{\xi_{i_1} = x\}) \wedge \Pi(\{\xi_{i_n} = y\}) \wedge \bigwedge_{j=2}^{n-1} \Pi(\{\xi_{i_j} = t_j\})\right)$$
$$= \bigvee_{\Gamma_1}\left(Q_\mu(x; a, b) \wedge Q_\mu(y; a, b) \wedge \bigwedge_{j=2}^{n-1} Q_\mu(t_j; a, b)\right)$$
$$= Q_\mu(x; a, b) \wedge Q_\mu(y; a, b)$$
$$= \Pi(\{m = x\}) \wedge \Pi(\{M = y\})$$

which completes the proof. □

**Proof of Theorem 3.1.** (a) The result that $\bar{X}$ has the same distribution function as $\xi_1$ is a consequence of Theorem 1 in [4]. Hence, $\Pi(\{\bar{X} = x\}) = Q_\mu(x; a, b)$ holds. This implies $\Pi(\{\bar{X} = a\}) = 1$ and $\bar{X}$ is a consistent estimator of the location parameter $a$. Let $x^{(1)}, x^{(2)}, \dots, x^{(n)}$ $(x^{(1)} \leqslant x^{(2)} \leqslant \cdots \leqslant x^{(n)})$ be the ordered values of the sample; we have

$$L(x_1, x_2, \dots, x_n, a) = \bigwedge_{j=1}^{n} Q_\mu(x_j; a, b) = Q_\mu(x^{(1)}; a, b) \wedge Q_\mu(x^{(n)}; a, b),$$

where $b$ is supposed to be known. It is easy to prove that $L$ attains maximum at $(x^{(1)} + x^{(n)})/2$ when $b$ is known. This results in the validity of part (a).

(b) Let $(x_1, x_2, \dots, x_n) \in R^n$, $\min(x_1, x_2, \dots, x_n) = \alpha - c\beta$ and $\mathrm{Max}(x_1, x_2, \dots, x_n) = \alpha + c\beta$. We have

$$\Pi(\{\xi_1 = x_1, \dots, \xi_n = x_n, \bar{X} = \alpha, S = \beta\})$$
$$= \Pi(\{\xi_1 = x_1, \dots, \xi_n = x_n, M = \alpha + c\beta, m = \alpha - c\beta\})$$
$$= Q_\mu(\alpha + c\beta; a, b) \wedge Q_\mu(\alpha - c\beta; a, b) \wedge \bigwedge_{j=2}^{n-1} Q_\mu(x^{(j)}; a, b)$$
$$= Q_\mu(\alpha + c\beta; a, b) \wedge Q_\mu(\alpha - c\beta; a, b)$$

(by monotonicity)

$$= \Pi(\{M = \alpha + c\beta\}) \wedge \Pi(\{m = \alpha - c\beta\})$$

(by Lemma 3.3)

$$= \Pi(\{M = \alpha + c\beta, m = \alpha - c\beta\})$$

(by Lemma 3.3)

$$= \Pi(\{\bar{X} = \alpha, S = \beta\}).$$

Thus $(\bar{X}, S)$ is sufficient with respect to $L(\mu)$. The proof is completed. □

## 4. Summary and conclusions

This paper is concerned with finding a solution to the Maxmin $\mu/E$ estimation for the family of fuzzy numbers with two parameters, location and scale, and investigating its characteristic properties. We gave a solution when the family is generated by a 0-symmetric fuzzy number. The solution could be regarded as an estimator whose sufficiency, consistency and maximum likelihood were shown.

The discussion was restricted to fuzzy number $\mu$ which is always 0-symmetric. Non-symmetric case remains to be studied further.

## References


[1] H. Dishkant, About membership functions estimation, *Fuzzy Sets and Systems* **5** (1981) 141–147.

[2] D. Dubois and H. Prade, *Fuzzy Sets and Systems: Theory and Applications* (Academic Press, New York, 1980).

[3] D. Dubois and H. Prade, *Possibility Theory* (Plenum Press, New York, 1988).

[4] W. Näther, On possibilistic inference, *Fuzzy Sets and Systems* **36** (1990) 327–337.

[5] W. Näther and M. Albrecht, Fuzzy model fitting based on the truth of the model, *Freiberger Forschungshefte* D 187 (1987).

[6] Wang Xizhao and Ha Minghu, Possibilistic statistics, *Proc. Sino–Japan Joint Meeting on Fuzzy Sets and Systems* (International Academic Publishers, 1990).

[7] Wang Xizhao and Ha Minghu, Fuzzy linear regression analysis, *Fuzzy Sets and Systems* **52** (1992) 179–188.

[8] W. Zhenyuan and L. Shoumei, Fuzzy linear regression analysis of fuzzy-valued variables, *Fuzzy Sets and Systems* **36** (1990) 125–136.