Cramér-Rao Lower Bound for Fuzzy-Valued Random Variables

Hamzeh Torabi Department of Statistics, Yazd University, Yazd, Iran

Abstract: In some point estimation problems, we may confront imprecise (fuzzy) concepts. One important case is a situation where all observations are fuzzy rather than crisp. In this paper, using fuzzy set theory, we define a fuzzy-valued random variable, a fuzzy unbiased estimator, a fuzzy exponential family, and then we state and prove a Cramér-Rao lower bound for such fuzzy estimators. Finally, we give some examples.

Keywords: Unbiased Estimator, Fuzzy Exponential Family, Cauchy-Schwartz Inequality.

1 Introduction

Point estimation in the traditional statistical inference is based on crispness of data, random variables, and so on. As there are many difficult situations in which the above assumptions are rather unrealistic, there have been some attempts to analyze theses situations with fuzzy set theory proposed by Zadeh (1965).

Fuzzy set theory is a powerful and known tool for formulation and analysis of imprecise and subjective situations where exact analysis is either difficult or impossible.

Some methods in descriptive statistics with vague data and some aspects of statistical inference is proposed in Kruse and Meyer (1987). Fuzzy random variables were introduced by Kwakernaak (1978), Puri and Ralescu (1986) as a generalization of compact random sets, Kruse and Meyer (1987) and were developed by others such as Juninig and Wang (1989), Ralescu (1995), López-Díaz and Gil (1997), M. López-Díaz (1998), and Liu (2004). Some aspects of point estimation problems with fuzzy data are discussed in Yao and Hwang (1996), Buckley (1985), Coral and Gil (1984), Gertner and Zhu (1996), Gil, Corral, and Gil (1985), Kruse (1984), Kruse and Meyer (1987), and Okuda (1987). For more details about ordinary point estimation problems (with crisp data), see Casella and Berger (2002), Freund (1992), Hogg and Craig (1995), Lehmann and Casella (1998), Mood, Graybill, and Boes (1974), and Shao (1998).

In this paper, because of our main purpose, statistical inference about a parametric population with fuzzy data, we only consider and discuss fuzzy random variables which associate with an ordinary random variable (fuzzy-valued random variable).

This paper is organized in the following way. In Section 2, we provide some definitions and preliminaries. A fuzzy unbiased estimator is defined in Section 3. The fuzzy exponential family is introduced in Section 4 and the Cramér-Rao lower bound is given in Section 5. Finally some applied examples are given in Section 6.

2 Preliminaries

Let (Ω, \mathcal{F}, P) be a probability space. A random variable (henceforth RV) X is a measurable function from (Ω, \mathcal{F}, P) to (R, \mathcal{B}, P_X) , where P_X is the probability measure induced by X and is called the distribution of the RV X, i.e.,

$$P_X(A) = P(X \in A) = \int_{X \in A} dP, \qquad A \in \mathcal{B}.$$

Using "the change of variable rule", (see e.g. Billingsley, 1995, pp. 215-216 or Shao, 1998, p. 13), we have

$$P_X(A) = \int_A dPoX^{-1}(x) = \int_A dP_X(x) , \qquad A \in \mathcal{B} .$$

If P_X is dominated by a σ -finite measure ν , i.e., $P_X \ll \nu$, then using the Radon-Nikodym theorem, (see for e.g. Billingsley, 1995, pp. 422-423 or Shao, 1998, p. 14), we have

$$P_X(A) = \int_A f(x) d\nu(x) \,,$$

where f(x) is the Radon-Nikodym derivative of P_X with respect to ν and is called the probability density function (henceforth PDF) of X with respect to ν .

In statistical texts, the measure ν usually is a "counting measure" or a "Lebesgue measure". Hence $P_X(A)$ is calculated as $\sum_{x \in A} f(x)$ or $\int_A f(x) dx$, respectively. Let $\mathcal{X} = \{x \in \mathbb{R} | f(x) > 0\}$. The set \mathcal{X} is usually called "support" or "sample space" of X.

Definition 2.1. $\mathbf{X} = (X_1, \dots, X_n)$ is a random sample of size *n* from a population with PDF f(x), if the X_i 's are independent and their PDF is f(x) (i.e. the X_i 's are identically distributed). In this case we have

$$f(\mathbf{x}) = f(x_1) \cdots f(x_n), \quad \forall x_i \in \mathbb{R},$$

where $\mathbf{x} = (x_1, \dots, x_n)$ is an observed value of **X**.

We now present two definitions from Casals et al. (1986), but in a slightly different way.

Definition 2.2. A fuzzy sample space $\hat{\mathcal{X}}$ is a fuzzy partition (Ruspini partition) of \mathcal{X} , i.e., a set of fuzzy subsets of \mathcal{X} whose membership functions are Borel measurable and satisfy the orthogonality constraint $\sum_{\tilde{x}\in\tilde{\mathcal{X}}} \mu_{\tilde{x}}(x) = 1$ for each $x \in \mathcal{X}$.

Definition 2.3. A fuzzy-valued random sample (henceforth FVRS) $\tilde{\mathbf{X}} = (\tilde{X}_1, \dots, \tilde{X}_n)$ of size *n* associated with PDF f(x) is a measurable function from Ω to $\tilde{\mathcal{X}}^n$, whose PDF is given by

$$\tilde{f}(\tilde{x}_1,\ldots,\tilde{x}_n)=\tilde{P}(\tilde{\mathbf{X}}=\tilde{\mathbf{x}})=\int_{\mathcal{X}^n}\prod_{i=1}^n\mu_{\tilde{x}_i}(x_i)f(x_i)d\nu(x_i)\,.$$

The above definition accords to Zadeh (1968). Note that using Fubini's theorem (see Billingsley, 1995, pp. 233-234), we obtain independency of \tilde{X}_i 's, i.e.,

$$\tilde{f}(\tilde{x}_1,\ldots,\tilde{x}_n) = \tilde{f}(\tilde{x}_1)\cdots\tilde{f}(\tilde{x}_n), \qquad \forall \tilde{x}_i \in \tilde{\mathcal{X}},$$

where

$$\tilde{f}(\tilde{x}_i) = \int_{\mathcal{X}} \mu_{\tilde{x}_i}(x_i) f(x_i) d\nu(x_i) \,,$$

and $\tilde{f}(\tilde{x}_i)$ is the PDF of the fuzzy-valued random variable (henceforth FVRV) \tilde{X}_i for each i = 1, ..., n. The $\tilde{f}(\tilde{x}_i)$ forms a PDF on $\tilde{\mathcal{X}}$, because by orthogonality of the $\mu_{\tilde{x}_i}$'s we have

$$\sum_{\tilde{x}_i \in \tilde{\mathcal{X}}} \tilde{f}(\tilde{x}_i) = \sum_{\tilde{x}_i \in \tilde{\mathcal{X}}} \int_{\mathcal{X}} \mu_{\tilde{x}_i}(x_i) f(x_i) d\nu(x_i)$$
$$= \int_{\mathcal{X}} f(x_i) \sum_{\tilde{x}_i \in \tilde{\mathcal{X}}} \mu_{\tilde{x}_i}(x_i) d\nu(x_i)$$
$$= \int_{\mathcal{X}} f(x_i) d\nu(x_i) = 1.$$

Theorem 2.1. If g is a measurable function from $\tilde{\mathcal{X}}^n$ to \mathbb{R} , then $Y = g(\tilde{\mathbf{X}})$ is an ordinary random variable.

Proof. The FVRS $\tilde{\mathbf{X}}$ is a measurable function from Ω to $\tilde{\mathcal{X}}^n$ and g is a measurable function from $\tilde{\mathcal{X}}^n$ to \mathbb{R} . Hence $g(\tilde{\mathbf{X}}(\omega)) = go\tilde{\mathbf{X}}(\omega)$ is a composition of two measurable functions, and therefore it is measurable from Ω to \mathbb{R} (see Billingsley, 1995, p. 182).

Note that using this theorem, we can define and use all related concepts of ordinary random variables, such as expectation, variance, etc.

Theorem 2.2. Let $\tilde{\mathbf{X}}$ be a FVRS with fuzzy sample space $\tilde{\mathcal{X}}^n$ and g a measurable function from $\tilde{\mathcal{X}}^n$ to \mathbb{R} . The expectation of $g(\tilde{\mathbf{X}})$ is calculated by

$$\mathbf{E}\left[g(\tilde{\mathbf{X}})\right] = \sum_{\tilde{\mathbf{x}} \in \tilde{\mathcal{X}}^n} g(\tilde{\mathbf{x}}) \tilde{f}(\tilde{\mathbf{x}}) \,.$$

Proof. Using the change of variable rule and the Radon-Nikodym theorem, we have

$$E\left[g(\tilde{\mathbf{X}})\right] = \int_{\Omega} g(\tilde{\mathbf{X}}(\omega)) dP(\omega)$$

=
$$\int_{\tilde{\mathcal{X}}^n} g(\tilde{\mathbf{x}}) dPo\tilde{\mathbf{X}}^{-1}(\tilde{\mathbf{x}})$$

=
$$\sum_{\tilde{\mathbf{x}} \in \tilde{\mathcal{X}}^n} g(\tilde{\mathbf{x}}) \tilde{f}(\tilde{\mathbf{x}}) .$$

For further details about the properties of ordinary RV's and their moments see Ash and Doleans-Dade (2000), Billingsley (1995), Chung (2000), Feller (1968), Ross (2002), and Shao (1998).

In this paper we suppose that the PDF of the population is known but has an unknown parameter $\theta \in \Theta$. In this case, we index \tilde{f} by θ and write $\tilde{f}(\tilde{\mathbf{x}}; \theta)$.

Example 2.1. Let X be a Bernoulli variable with parameter θ , i.e.,

$$f(x; \theta) = \theta^x (1 - \theta)^{1-x}, \qquad x = \in \{0, 1\}, \ 0 < \theta < 1.$$

We have $\mathcal{X} = \{0, 1\}$. Let \tilde{x}_1 and \tilde{x}_2 be two fuzzy subsets of \mathcal{X} with membership functions

$$\mu_{\tilde{x}_1}(x) = \begin{cases} 0.9, & x = 0\\ 0.1, & x = 1 \end{cases} \quad \text{and} \quad \mu_{\tilde{x}_2}(x) = \begin{cases} 0.1, & x = 0\\ 0.9, & x = 1 \end{cases}$$

Note that \tilde{x}_1 and \tilde{x}_2 can be regarded as "approximately zero" and "approximately one", respectively. In this example, the support of \tilde{X} is $\tilde{\mathcal{X}} = {\tilde{x}_1, \tilde{x}_2}$ and using Definition 2.3 for n = 1, the PDF of \tilde{X} is

$$\tilde{f}(\tilde{x};\theta) = \sum_{\mathcal{X}} \mu_{\tilde{x}_1}(x) f(x;\theta) = \begin{cases} 0.9(1-\theta) + 0.1\theta = 0.9 - 0.8\theta, & \tilde{x} = \tilde{x}_1 \\ 0.1(1-\theta) + 0.9\theta = 0.1 + 0.8\theta, & \tilde{x} = \tilde{x}_2. \end{cases}$$

Now, e.g., consider

$$Y = \begin{cases} 0.1 \, , \ \tilde{x} = \tilde{x}_1 \\ 0.9 \, , \ \tilde{x} = \tilde{x}_2 \, . \end{cases}$$

Note that Y is a measurable function from \hat{X} to \mathbb{R} and therefore an ordinary RV. In the following we explain how to calculate the mean and variance of Y. The PDF of Y is

$$f_Y(y) = \begin{cases} 0.9 - 0.8\theta, & y = 0.1\\ 0.1 + 0.8\theta, & y = 0.9. \end{cases}$$

Therefore, the mean and the variance of Y are calculated as

 $E[Y] = 0.1(0.9 - 0.8\theta) + 0.9(0.1 + 0.8\theta) = 0.18 + 0.64\theta$

and

$$E[Y^2] = 0.01(0.9 - 0.8\theta) + 0.81(0.1 + 0.8\theta) = 0.09 + 0.64\theta$$

var[Y] = 0.09 + 0.64\theta - (0.18 + 0.64\theta)^2 = 0.0576 + 0.4096\theta - 0.4096\theta^2.

3 Fuzzy Unbiased Estimator

Let $\tilde{\mathbf{X}} = (\tilde{X}_1, \dots, \tilde{X}_n)$ be a FVRS with the common PDF $\tilde{f}(\tilde{\mathbf{x}}; \theta)$, where $\theta \in \Theta \subset \mathbb{R}$. Assume that $\tilde{\mathbf{x}} = (\tilde{x}_1, \dots, \tilde{x}_n)$ is an observation of $\tilde{\mathbf{X}}$, i.e., $\tilde{\mathbf{x}} \in \tilde{\mathcal{X}}^n$.

Definition 3.1. Every real-valued measurable function of \mathbf{X} , say $T(\mathbf{X})$, which is independent of the unknown parameter is a fuzzy estimator for θ . We say $T(\tilde{\mathbf{x}})$ is a fuzzy estimate for θ .

Definition 3.2. A fuzzy estimator $T(\mathbf{X})$ is a fuzzy unbiased estimator for $\gamma(\theta)$, if

$$\mathbf{E}[T(\mathbf{X})] = \gamma(\theta) \,, \qquad \forall \theta \in \Theta \,.$$

Definition 3.3. If there exists an unbiased estimator $T(\mathbf{X})$ of $\gamma(\theta)$, then $\gamma(\theta)$ is called an estimable parameter (or U-estimable).

Definition 3.4. A fuzzy unbiased estimator $T(\mathbf{X})$ of $\gamma(\theta)$ is called the uniformly minimum variance unbiased estimator (henceforth UMVUE), if $var[T(\tilde{\mathbf{X}})] \leq var[T^*(\tilde{\mathbf{X}})]$ for any $\theta \in \Theta$ and any other fuzzy unbiased estimator $T^*(\tilde{\mathbf{X}})$ of $\gamma(\theta)$. **Example 3.1.** Let $\tilde{X}_1, \ldots, \tilde{X}_n$ be a sample from the distribution of the FVRV \tilde{X} of Example 2.1 with substituting \tilde{x}_1 and \tilde{x}_2 with \tilde{x}_I and \tilde{x}_{II} , respectively. We saw that

$$\tilde{f}(\tilde{x};\theta) = \begin{cases} 0.9 - 0.8\theta, \ \tilde{x} = \tilde{x}_I\\ 0.1 + 0.8\theta, \ \tilde{x} = \tilde{x}_{II}. \end{cases}$$

Using Definition 3.2 in case n = 1, the fuzzy estimator $T(\tilde{X})$ is a fuzzy unbiased estimator for θ , if

$$\mathbf{E}[T(\tilde{X})] = T(\tilde{x}_I)\tilde{f}(\tilde{x}_I;\theta) + T(\tilde{x}_{II})\tilde{f}(\tilde{x}_{II};\theta) = \theta, \qquad \forall \theta \in \Theta = (0,1).$$

Hence, we must solve the system

$$0.9T(\tilde{x}_I) + 0.1T(\tilde{x}_{II}) = 0$$

$$0.1T(\tilde{x}_I) + 0.9T(\tilde{x}_{II}) = 1.$$

But its unique solution is

$$T(\tilde{X}) = \begin{cases} -1/8 , \ \tilde{X} = \tilde{x}_I \\ 9/8 , \ \tilde{X} = \tilde{x}_{II} \end{cases}$$

Thus, for n = 1, $T(\tilde{X})$ is a fuzzy unbiased estimator. Note that with crisp data ($\mathcal{X} = \{0, 1\}$), the unique unbiased estimator is X. Hence, this result approximately coincides with the classical one. But the main disadvantage of this fuzzy estimator is that it is not "range preserve", i.e., $T(\tilde{X}) \notin \Theta$. Of course this problem occurs in some classical cases, e.g., let $\Theta = (0, 1)$ then neither 0 nor 1 belongs to Θ . Now consider fuzzy estimator $S(\tilde{X}) = \sum_{i=1}^{n} T(\tilde{X}_i)/n$. The fuzzy estimator $S(\tilde{X})$ is the sample mean of -1/8 and 9/8, therefore it slightly obviates this problem. Using Theorem 2.2, $S(\tilde{X})$ is a fuzzy unbiased estimator for θ .

The variance of T(X) is

$$\operatorname{var}[T(\tilde{X})] = \frac{1}{64}(0.9 - 0.8\theta) + \frac{81}{64}(0.1 + 0.8\theta) - \theta^2 = \frac{9}{64} + \theta - \theta^2.$$

Because of the independency of the \tilde{X}_i 's, and thus by the independency of the $T(\tilde{X}_i)$'s, we have $var[S(\tilde{\mathbf{X}})] = (9/64 + \theta - \theta^2)/n$.

Example 3.2. Let X_1, \ldots, X_n be a sample from the Beta $(\theta, 1)$ population, i.e.,

$$f(x;\theta) = \theta x^{\theta-1}, \qquad 0 < x < 1, \ \theta > 0.$$

Let \tilde{X} take two fuzzy subsets \tilde{x}_I and \tilde{x}_{II} , whose membership functions are

$$\mu_{\tilde{x}_I} = x, \quad 0 < x < 1, \qquad \mu_{\tilde{x}_{II}} = 1 - x, \quad 0 < x < 1.$$

Note that \tilde{x}_I and \tilde{x}_{II} can be regarded as "near to one" and "near to zero", respectively. We have

$$\tilde{f}(\tilde{x};\theta) = \int_0^1 \mu_{\tilde{x}}(x)\theta x^{\theta-1} = \begin{cases} \frac{\theta}{\theta+1}, & \tilde{x} = \tilde{x}_I\\ 1 - \frac{\theta}{\theta+1}, & \tilde{x} = \tilde{x}_{II} \end{cases}$$

For $n = 1, T(\tilde{X})$ is a fuzzy unbiased estimator of θ , if

$$\mathbf{E}[T(\tilde{X})] = T(\tilde{x}_I)\tilde{f}(\tilde{x}_I;\theta) + T(\tilde{x}_{II})\tilde{f}(\tilde{x}_{II};\theta) = \theta, \qquad \forall \theta \in \Theta = (0,+\infty)$$

Therefore, we must have $\theta^2 + \theta(1 - T(\tilde{x}_I)) - T(\tilde{x}_I) = 0$, $\forall \theta > 0$. But this is impossible. Hence, in this case there does not exist any fuzzy unbiased estimator for θ , i.e., θ is not an estimable parameter.

Now, suppose that we want to find a fuzzy unbiased estimator for $\theta/(\theta + 1)$. We can use the fuzzy estimator $I_{\{\tilde{x}_I\}}(\tilde{X})$; since $\mathbb{E}[I_{\{\tilde{x}_I\}}(\tilde{X})] = \tilde{P}(\tilde{X} = \tilde{x}_I) = \theta/(\theta + 1)$, $S(\tilde{\mathbf{X}}) = \sum_{i=1}^n I_{\{\tilde{x}_I\}}(\tilde{X}_i)/n$ is a fuzzy unbiased estimator for θ . Because $\operatorname{var}[I_{\{\tilde{x}_I\}}(\tilde{X})] = \theta/(\theta + 1)^2$ and because the \tilde{X}_i 's are independent, we have

$$\operatorname{var}[S(\tilde{\mathbf{X}})] = \frac{\theta}{n(\theta+1)^2}$$

4 Fuzzy Exponential Family

In this section we introduce the fuzzy exponential family.

Definition 4.1. A fuzzy parametric family $\{\tilde{f}(\tilde{x};\theta), \theta \in \Theta\}$ of fuzzy density functions is said to form an *s*-dimensional fuzzy exponential family, if

$$\tilde{f}(\tilde{x};\theta) = a(\theta)b(\tilde{x})\exp\left\{\sum_{i=1}^{s}c_{i}(\theta)d_{i}(\tilde{x})\right\},\$$

where $a(\theta)$ and $c_i(\theta)$, i = 1, ..., s, are real-valued functions of the parameter and $b(\tilde{x})$ and $d_i(\tilde{x})$ are Borel measurable functions of $\tilde{x}, \tilde{x} \in \tilde{\mathcal{X}}$. Also $a(\theta) > 0$ for all $\theta \in \Theta$ and $b(\tilde{x}) > 0$ for all $\tilde{x} \in \tilde{\mathcal{X}}$.

Theorem 4.1. Let \tilde{X} be a FVRV (associated with the RV X) with PDF $\tilde{f}(\tilde{x};\theta), \theta \in \Theta$, and let its support be independent of θ and be a finite *m*-set of fuzzy observation, i.e., $\tilde{X} = {\tilde{x}_1, \ldots, \tilde{x}_m}$. Then the PDF $\tilde{f}(\tilde{x};\theta)$ belongs to the fuzzy exponential family. **Proof.** Using Definition 2.3, we have

$$\tilde{f}(\tilde{x};\theta) = \int_{\mathcal{X}} \mu_{\tilde{x}}(x) f(x;\theta) d\nu(x), \qquad \tilde{x} = \tilde{x}_1, \dots, \tilde{x}_m.$$

Let $p_i(\theta) = \int_{\mathcal{X}} \mu_{\tilde{x}_i}(x) f(x; \theta) d\nu(x)$. Hence,

$$\tilde{f}(\tilde{x};\theta) = \begin{cases} p_1(\theta), & \tilde{x} = \tilde{x}_1 \\ \vdots & \vdots \\ p_m(\theta), & \tilde{x} = \tilde{x}_m \end{cases}$$
$$= \left[\prod_{i=1}^{m-1} p_i(\theta)^{I_{\{\tilde{x}_i\}}(\tilde{x})}\right] \times p_m(\theta)^{1-\sum_{i=1}^{m-1} I_{\{\tilde{x}_i\}}(\tilde{x})}$$
$$= p_m(\theta) \exp\left\{\sum_{i=1}^{m-1} I_{\{\tilde{x}_i\}}(\tilde{x}) \left[\log(p_i(\theta)) - \log(p_m(\theta))\right]\right\}$$

Thus, $\tilde{f}(\tilde{x};\theta)$ belongs to the (m-1)-dimensional fuzzy exponential family with $a(\theta) = p_m(\theta), b(\tilde{x}) = 1, c_i(\theta) = \log(p_i(\theta)) - \log(p_m(\theta), \text{ and } d_i(\tilde{x}) = I_{\{\tilde{x}_i\}}(\tilde{x}).$

Note that in Theorem 4.1, the dimension is at most m - 1, because some of the p_i 's may be equal; see Example 6.3.

5 Cramér-Rao Lower Bound

In this section, we state and prove a new version of Cramér-Rao lower bound for FVRV's; see Lehmann and Casella (1998, p. 120) or Shao (1998, pp. 135-136) for ordinary RV's.

Theorem 5.1. (Cramér-Rao lower bound for FVRV's) Let $\tilde{\mathbf{X}}$ be a FVRS from the PDF $\tilde{f}(\tilde{\mathbf{x}}; \theta), \theta \in \Theta$, where Θ is the parameter space and be an open interval of the real line. If $\tilde{f}(\tilde{\mathbf{x}}; \theta)$ is differentiable as a function of θ and satisfies

$$\frac{\partial}{\partial \theta} \sum_{\tilde{\mathcal{X}}^n} h(\tilde{\mathbf{x}}) \tilde{f}(\tilde{\mathbf{x}}; \theta) = \sum_{\tilde{\mathcal{X}}^n} h(\tilde{\mathbf{x}}) \frac{\partial}{\partial \theta} \tilde{f}(\tilde{\mathbf{x}}; \theta)$$
(1)

for both, $h(\tilde{\mathbf{x}}) = 1$ and $h(\tilde{\mathbf{x}}) = T(\tilde{\mathbf{x}})$, and $\gamma(\theta) = E[T(\tilde{\mathbf{X}})]$ is a differentiable function of θ , then the variance of each fuzzy estimator of the form $T(\tilde{\mathbf{X}})$ satisfies

$$\operatorname{var}[T(\tilde{\mathbf{X}})] \geq \frac{[\gamma'(\theta)]^2}{\operatorname{E}\left[\frac{\partial}{\partial \theta} \log \tilde{f}(\tilde{\mathbf{X}}; \theta)\right]^2}, \qquad \forall \theta \in \Theta.$$

Proof. Let $Y = \partial \log \tilde{f}(\tilde{\mathbf{X}}; \theta) / \partial \theta$, then Y and $T(\tilde{\mathbf{X}})$ are ordinary RV's. Hence from the Cauchy-Schwartz inequality, we only need to show that $E[Y^2] = var[Y]$ and $\gamma'(\theta) = cov(T(\tilde{\mathbf{X}}), Y)$. But these two results are consequences of condition (1).

Note that in Theorem 5.1, if $\tilde{\mathcal{X}}$ is a finite fuzzy partition of \mathcal{X} , then the condition (1) is satisfied.

The right hand side of the inequality of Theorem 5.1, is called the Cramér-Rao lower bound (henceforth CRLB). If the variance of a fuzzy estimator equals the CRLB, we say that this fuzzy estimator (or its variance) attains the CRLB. In this case it is a UMVUE for its expectation. Note that the UMVUE is a fuzzy unbiased estimator with minimum variance, not necessarily a fuzzy unbiased estimator where its variance attains the CRLB. We will illustrate that this inequality is sometimes sharp, i.e., there is no unbiased estimator whose variance equals the CRLB, see Example 6.3 for more details.

We call $\mathbb{E}[\partial \log \tilde{f}(\tilde{\mathbf{X}}; \theta) / \partial \theta]^2$ the fuzzy information of $\tilde{\mathbf{X}}$ and denote it by $\tilde{I}_{\tilde{\mathbf{X}}}(\theta)$. Note that $\tilde{I}_{\tilde{\mathbf{X}}}(\theta)$ is the amount of information that $\tilde{\mathbf{X}}$ contain about the parameter θ .

Theorem 5.2. (Attainment) Suppose that the assumptions of Theorem 5.1 are satisfied. Then

$$\operatorname{var}[T(\tilde{\mathbf{X}})] = \frac{[\gamma'(\theta)]^2}{\operatorname{E}\left[\frac{\partial}{\partial \theta} \log \tilde{f}(\tilde{\mathbf{X}}; \theta)\right]^2}, \qquad \forall \theta \in \Theta,$$

if and only if there is a continuously differentiable function $c(\theta)$ such that

$$\tilde{f}(\tilde{\mathbf{x}};\theta) = a(\theta)b(\tilde{\mathbf{x}})e^{c(\theta)d(\tilde{\mathbf{x}})},$$

is a density function for suitably chosen $a(\theta)$ and $b(\tilde{\mathbf{x}})$, i.e., $\tilde{f}(\tilde{x};\theta)$ constitutes an exponential family. Moreover $d(\tilde{\mathbf{x}})$ and $\gamma(\theta)$ satisfy

$$d(\tilde{\mathbf{x}}) = \left[\frac{\gamma'(\theta)}{\tilde{I}_{\tilde{\mathbf{X}}}(\theta)}\right] \frac{\partial}{\partial \theta} \log \tilde{f}(\tilde{\mathbf{x}};\theta)) + \gamma(\theta), \qquad \gamma(\theta) = -\frac{a'(\theta)}{a(\theta)}c'(\theta), \tag{2}$$

and $\tilde{I}_{\tilde{\mathbf{X}}}(\theta) = c'(\theta)\gamma'(\theta)$.

Note that the function $d(\tilde{\mathbf{x}})$ specified in (2) may depend on θ . In such a case, $d(\tilde{\mathbf{X}})$ is not a fuzzy estimator, and there is no estimator that attains the Cramér-Rao lower bound. **Proof.** See Shao (1998) for the proof of ordinary RV's. Using Theorem 2.2, substitute \mathbf{X} by $\tilde{\mathbf{X}}$, for completing the proof.

Note that if $T(\tilde{\mathbf{X}})$ is an unbiased estimator for θ , then under the assumptions of Theorem 5.1, we have

$$\operatorname{var}[T(\tilde{\mathbf{X}})] \geq \frac{1}{\operatorname{E}\left[\frac{\partial}{\partial \theta} \log \tilde{f}(\tilde{\mathbf{X}}; \theta)\right]^{2}}.$$

Theorem 5.3. If in Theorem 5.1, $\tilde{\mathbf{X}}$ is a FVRS from the PDF $\tilde{f}(\tilde{x}; \theta)$, then

$$\tilde{I}_{\tilde{\mathbf{X}}}(\theta) = n\tilde{I}_{\tilde{X}}(\theta)$$

Proof. We have

$$\tilde{I}_{\tilde{\mathbf{X}}}(\theta) = E\left[\sum_{i=1}^{n} \frac{\partial}{\partial \theta} \log \tilde{f}(\tilde{X}_{i};\theta)\right]^{2}$$
$$= \sum_{i=1}^{n} \sum_{j=1}^{n} E\left[\frac{\partial}{\partial \theta} \log \tilde{f}(\tilde{X}_{i};\theta) \times \frac{\partial}{\partial \theta} \log \tilde{f}(\tilde{X}_{j};\theta)\right].$$

But by independency of \tilde{X}_i 's, we get

$$E\left[\left(\frac{\partial}{\partial\theta}\log\tilde{f}(\tilde{X}_i;\theta)\right)\times\left(\frac{\partial}{\partial\theta}\log\tilde{f}(\tilde{X}_j;\theta)\right)\right] = \begin{cases} I_{\tilde{X}_i}(\theta), & i=j\\ 0, & i\neq j \end{cases}$$

Theorem 5.4. Suppose that $\tilde{\mathbf{X}}$ has PDF $\tilde{f}(\tilde{\mathbf{x}}; \theta)$, which is twice differentiable in θ and condition (1) holds with $h(\tilde{\mathbf{x}}) = 1$ and $\tilde{f}(\tilde{x}; \theta)$ replaced by $\partial \tilde{f}(\tilde{\mathbf{x}}; \theta)/\partial \theta$, then

$$I_{\tilde{\mathbf{X}}}(\theta) = -\mathbf{E}\left[\frac{\partial^2}{\partial\theta^2}\log \tilde{f}(\tilde{\mathbf{X}};\theta)\right].$$

Proof. The theorem follows from the equality

$$\frac{\partial^2}{\partial\theta^2}\log\tilde{f}(\tilde{\mathbf{X}};\theta) = \frac{1}{\tilde{f}(\tilde{\mathbf{X}};\theta)} \left[\frac{\partial^2}{\partial\theta^2}\tilde{f}(\tilde{\mathbf{X}};\theta)\right] - \left(\frac{\partial}{\partial\theta}\log\tilde{f}(\tilde{\mathbf{X}};\theta)\right)^2$$

6 Some Examples

Example 6.1. Remember Example 3.1. It was mentioned that the fuzzy estimator $S(\tilde{\mathbf{X}}) = \sum_{i=1}^{n} T(\tilde{X}_i)/n$ is a fuzzy unbiased estimator for θ with $var[S(\tilde{\mathbf{x}})] = (9/64 + \theta - \theta^2)/n$, where

$$T(\tilde{X}_{i}) = \begin{cases} -1/8, \ X_{i} = \tilde{x}_{I} \\ 9/8, \ \tilde{X}_{i} = \tilde{x}_{II}. \end{cases}$$

We want to prove that this fuzzy estimator attains the CRLB, i.e.,

$$\operatorname{var}[S(\tilde{\mathbf{X}})] = \frac{1}{n \operatorname{E}\left[\frac{\partial}{\partial \theta} \log \tilde{f}(\tilde{X}; \theta)\right]^2}$$

We offer two proofs:

Proof (1): By Theorem 5.2, this family belongs to the fuzzy exponential family, because

$$\tilde{f}(\tilde{x};\theta) = (0.1+0.8\theta) \exp\left\{I_{\{\tilde{x}_I\}}(\tilde{x})\log\frac{0.9-0.8\theta}{0.1+0.8\theta}\right\}, \qquad \tilde{x} = \tilde{x}_I, \tilde{x}_{II}$$

Take $a(\theta) = 0.1 + 0.8\theta$, $b(\tilde{x}) = 1$, $c(\theta) = \log(0.9 - 0.8\theta)/(0.1 + 0.8\theta)$, and $d(\tilde{x}) = I_{\{\tilde{x}_I\}}(\tilde{x})$. Then using Theorem 5.2, the variance of $T^*(\tilde{X}) = I_{\{\tilde{x}_I\}}(\tilde{X})$ attains the CRLB for the parameter $E[T^*(\tilde{X})] = 0.9 - 0.8\theta$. Hence, $(0.9 - T^*(\tilde{X}))/0.8$ is an unbiased estimator of θ ; but $T(\tilde{X}) = (0.9 - T^*(\tilde{X}))/0.8$. Because of the linear relationship between $T^*(\tilde{X})$ and $T(\tilde{X})$, using Theorem 5.2, the claim is proved.

Proof (2): (direct) We have

$$\frac{\partial}{\partial \theta} \log \tilde{f}(\tilde{X}; \theta) = \begin{cases} \frac{\partial}{\partial \theta} \log(0.9 - 0.8\theta) = \frac{-0.8}{0.9 - 0.8}, & \tilde{X} = \tilde{x}_I \\ \frac{\partial}{\partial \theta} \log(0.1 + 0.8\theta) = \frac{0.8}{0.1 + 0.8}, & \tilde{X} = \tilde{x}_{II}. \end{cases}$$

Hence

$$\operatorname{E}\left[\frac{\partial}{\partial\theta}\log\tilde{f}(\tilde{X};\theta)\right]^{2} = \frac{0.64}{0.9 - 0.8\theta} + \frac{0.64}{0.1 + 0.8\theta}$$

Therefore,

$$\text{CRLB} = \frac{1}{n\left(\frac{0.64}{0.9 - 0.8\theta} + \frac{0.64}{0.1 + 0.8\theta}\right)} = \frac{9}{64} + \theta - \frac{\theta^2}{n}$$

Thus, the fuzzy estimator $S(\tilde{\mathbf{X}})$ attains the CRLB, i.e., $S(\tilde{\mathbf{X}})$ is the UMVUE of θ .

Example 6.2. Consider again Example 3.2. We have seen that $S(\tilde{\mathbf{X}}) = \sum_{i=1}^{n} T(\tilde{X}_i)/n$ is a fuzzy unbiased estimator with $\operatorname{var}[S(\tilde{\mathbf{X}})] = \theta/(n(\theta+1)^2)$, where

$$T(\tilde{X}_i) = I_{\{\tilde{x}_I\}}(\tilde{X}_i) \,.$$

Let $\lambda = \theta/(\theta + 1)$. We want to prove that this fuzzy estimator attains the CRLB, i.e.,

$$\operatorname{var}[S(\tilde{\mathbf{X}})] = \frac{1}{\operatorname{E}\left[\frac{\partial}{\partial\lambda}\log\tilde{f}(\tilde{X};\theta)\right]^{2}}.$$

We offer two proofs:

Proof (1): By Theorem 5.2 this family belongs to the fuzzy exponential family, because

$$\tilde{f}(\tilde{x};\theta) = \frac{1}{\theta+1} \exp\{I_{\{\tilde{x}_I\}}(\tilde{x})\log(\theta)\}, \qquad \tilde{x} = \tilde{x}_I, \tilde{x}_{II}.$$

Take $a(\theta) = 1/(\theta + 1)$, $b(\tilde{x}) = 1$, $c(\theta) = \log(\theta)$, and $d(\tilde{x}) = I_{\{\tilde{x}_I\}}(\tilde{x})$. Therefore, using Theorem 5.2, the variance of the fuzzy estimator $T(\tilde{X}) = I_{\{\tilde{x}_I\}}(\tilde{x})$ attains the CRLB for the parameter $E[T(\tilde{X})] = \theta/(\theta + 1) = \lambda$. Proof (2) (direct) We have

$$\frac{\partial}{\partial\lambda}\log\tilde{f}(\tilde{X};\theta) = \begin{cases} \frac{\partial}{\partial\lambda}\log\lambda = \frac{1}{\lambda}, & \tilde{X} = \tilde{x}_I\\ \frac{\partial}{\partial\lambda}\log(1-\lambda) = -\frac{1}{1-\lambda}, & \tilde{X} = \tilde{x}_{II}. \end{cases}$$

Hence

$$\operatorname{E}\left[\frac{\partial}{\partial\theta}\log\tilde{f}(\tilde{X};\theta)\right]^{2} = \frac{1}{\lambda} + \frac{1}{1-\lambda} = \frac{1}{\lambda(1-\lambda)}$$

Therefore,

$$\mathbf{CRLB} = \frac{\lambda(1-\lambda)}{n}$$

Thus, the fuzzy estimator $S(\tilde{\mathbf{X}})$ attains the CRLB, i.e., $S(\tilde{\mathbf{X}})$ is the UMVUE of $\lambda =$ $\theta/(\theta+1).$

Example 6.3. Let $X \sim N(0, \theta^2), \theta > 0$. We have $\mathcal{X} = \mathbb{R}$. Assume that its observation is as three fuzzy data $\tilde{x}_I, \tilde{x}_{II}$, and \tilde{x}_{III} , whose membership functions are

$$\mu_{\tilde{x}_{I}}(x) = \begin{cases} 1 - \exp\{-x^{2}/2\}, & x < 0\\ 0, & x \ge 0, \end{cases} \qquad \mu_{\tilde{x}_{II}}(x) = \exp\{-x^{2}/2\}, & x \in \mathbb{R}, \end{cases}$$

and

$$\mu_{\tilde{x}_{III}}(x) = \begin{cases} 0, & x < 0\\ 1 - \exp\{-x^2/2\}, & x \ge 0 \end{cases}$$

Note that \tilde{x}_I , \tilde{x}_{II} , and \tilde{x}_{III} can be stated as "small and negative", "more or less zero", and "large and positive", respectively.

We have

$$\tilde{f}(\tilde{x};\theta) = \begin{cases} 1/2 - 1/[2\sqrt{\theta^2 + 1}], \quad \tilde{x} = \tilde{x}_I \\ 1/\sqrt{\theta^2 + 1}, \quad \tilde{x} = \tilde{x}_{II} \\ 1/2 - 1/[2\sqrt{\theta^2 + 1}], \quad \tilde{x} = \tilde{x}_{III} \end{cases}$$
$$= \left(1/\sqrt{\theta^2 + 1}\right)^{I_{\{\tilde{x}_{II}\}}(\tilde{x})} \left(1/2 - 1/(2\sqrt{\theta^2 + 1})\right)^{1 - I_{\{\tilde{x}_{II}\}}(\tilde{x})} \\= \frac{1}{2} \left(1 - 1/\sqrt{\theta^2 + 1}\right) \exp\left\{I_{\{\tilde{x}_{II}\}}(\tilde{x})\log\frac{2}{\sqrt{\theta^2 + 1} - 1}\right\}.$$

Hence, the family belongs to the fuzzy exponential family with $a(\theta) = (1-1/\sqrt{\theta^2+1})/2$, $b(\tilde{x}) = 1, c(\theta) = \log(2/(\sqrt{\theta^2 + 1} - 1)), \text{ and } d(\tilde{x}) = I_{\{\tilde{x}_{II}\}}(\tilde{x}).$ Therefore, with a sample of size *n*, the variance of the fuzzy estimator $S(\tilde{\mathbf{X}}) = \sum_{i=1}^{n} I_{\{\tilde{x}_{II}\}}(\tilde{X}_i)/n$ attains the CRLB for its expectation $1/\sqrt{\theta^2 + 1}$, i.e., $T(\tilde{\mathbf{X}})$ is the UMVUE of $1/\sqrt{\theta^2 + 1}$.

Let $T(\tilde{X}) = I_{\{\tilde{x}_I\}}(\tilde{X})$. We have $E[T(\tilde{X})] = [1 - 1/\sqrt{\theta^2 + 1}]/2$. This fuzzy estimator does not attain the CRLB for its expectation, since

$$\operatorname{var}[T(\tilde{X})] = \frac{1}{2} \left[1 - \frac{1}{\sqrt{\theta^2 + 1}} \right] - \left[\frac{1}{2} \left(1 - \frac{1}{\sqrt{\theta^2 + 1}} \right) \right]^2 = \frac{\theta^2}{4(\theta^2 + 1)},$$

and

$$\frac{\partial}{\partial \theta} \log \tilde{f}(\tilde{X};\theta) = \begin{cases} \frac{\partial}{\partial \theta} \log((1/2)[1-1/\sqrt{\theta^2+1}]) = \frac{\theta}{(1+\theta^2)(-1+\sqrt{\theta^2+1})}, & \tilde{x} = \tilde{x}_I\\ \frac{\partial}{\partial \theta} \log((1/\sqrt{\theta^2+1})) = -\frac{\theta}{\theta^2+1}, & \tilde{x} = \tilde{x}_{II}\\ \frac{\partial}{\partial \theta} \log((1/2)[1-1/\sqrt{\theta^2+1}]) = \frac{\theta}{(1+\theta^2)(-1+\sqrt{\theta^2+1})}, & \tilde{x} = \tilde{x}_{III}. \end{cases}$$

Hence

$$\mathbf{E}\left[\frac{\partial}{\partial\lambda}\log\tilde{f}(\tilde{X};\theta)\right]^2 = \frac{\theta^2}{(\theta^2+1)^2 + (-1+\sqrt{\theta^2+1})}$$

But $\gamma'(\theta) = \theta/(2(\theta^2+1)^{3/2})$. Thus,

$$CRLB = \frac{-1 + \sqrt{\theta^2 + 1}}{4(\theta^2 + 1)^2},$$

which is less than $var[T(\tilde{X})]$ for all $\theta > 0$, because $\sqrt{\theta^2 + 1} < \theta^2 + 1$. Therefore, $T(\tilde{X})$ does not attain the CRLB.

References

- Ash, R. B., and Doleans-Dade, C. A. (2000). *Probability and Measure Theory* (2nd ed.). Academic Press.
- Billingsley, P. (1995). Probability and Measure (3rd ed.). John Wiley & Sons.
- Buckley, J. J. (1985). Fuzzy decision making with data: Applications to statistics. *Fuzzy Sets and Systems*, *16*, 139-147.
- Casals, M. R., Gil, M. R., and Gil, P. (1986). On the use of Zadeh's probabilistic definition for testing statistical hypotheses from fuzzy information. *Fuzzy Sets and Systems*, 20, 175-190.
- Casella, G., and Berger, R. L. (2002). Statistical Inference (2nd ed.). Duxbury Press.

Chung, K. L. (2000). A Course in Probability Theory (2nd ed.). Academic Press.

- Coral, N., and Gil, M. A. (1984). The minimum inaccuracy fuzzy estimation: An extension of the maximum likelihood principle. *Stochastica*, *6*, 63-81.
- Feller, W. (1968). An Introduction to Probability Theory and Its Applications, Volume I (3rd ed.). John Wiley & Sons.

Freund, J. E. (1992). Mathematical Statistics (5th ed.). Prentice Hall.

Gertner, G. Z., and Zhu, H. (1996). Bayesian estimation in forest surveys when samples or prior information are fuzzy. *Fuzzy Sets and Systems*, 77, 277-290.

- Gil, M. A., Corral, N., and Gil, P. (1985). The fuzzy decision problem: An approach to the point estimation problem with fuzzy information. *European Journal of Operation Research*, 22, 26-34.
- Hogg, R. V., and Craig, A. T. (1995). *Introduction to Mathematical Statistics* (5th ed.). Eaglewood Cliffs, NJ: Prentice-Hall.
- Juninig, O., and Wang, G. (1989). Fuzzy random variables and their probability characters. *Journal of Harbin Architectural and Civil Engineering Institute*, *12*, 1-10.
- Kruse, R. (1984). Statistical estimation with linguistic data. *Information Sciences*, *33*, 197-207.
- Kruse, R., and Meyer, K. D. (1987). *Statistics with Vague Data*. Dordrech, Netherlands: Reidel Pub. Comp.
- Kwakernaak, H. (1978). Fuzzy random variables: Definition and theorems. *Information Sciences*, *15*, 1-29.
- Lehmann, E. L., and Casella, G. (1998). Theory of Point Estimation. Springer-Verlag.
- Liu, B. (2004). Uncertainty Theory: An Introduction to Its Axiomatic Foundation. Heidelberg: Physica-Verlag.
- López-Díaz, M., and Gil, M. A. (1997). Constructive definitions of fuzzy random variables. *Statistics and Probability Letters*, *36*, 135-144.
- M. López-Díaz, M. A. G. (1998). Approximately integrable bounded fuzzy random variable in terms of the "generalized" Hausdorff metric. *Information Sciences*, 104, 279-291.
- Mood, A. M., Graybill, F. A., and Boes, D. C. (1974). *Introduction to the Theory of Statistics* (3rd ed.). McGraw-Hill.
- Okuda, T. (1987). A statistical treatment of fuzzy observations: Estimation problems. 51-55. (Preprints of the Second IFSA Congress)
- Puri, M. L., and Ralescu, D. A. (1986). Fuzzy random variables. *Journal of Mathematical Analysis and Applications*, 114, 409-422.
- Ralescu, D. A. (1995). Fuzzy random variable revisited. In *Proc. IFES' 95, Vol. 2* (p. 993-1000).
- Ross, S. M. (2002). A First Course in Probability (6th ed.). Prentice-Hall.
- Shao, J. (1998). Mathematical Statistics. New York: Springer-Verlag.
- Yao, J. S., and Hwang, C. M. (1996). Point estimation for the *n* sizes of random sample with one vague data. *Fuzzy Sets and Systems*, 80, 205-215.
- Zadeh, L. A. (1965). Fuzzy sets. Information and Control, 8, 338-353.
- Zadeh, L. A. (1968). Probability measure of fuzzy events. *Journal of Mathematical Analysis and Applications*, 23, 421-427.

Authors' address:

Hamzeh Torabi Department of Statistics Yazd University Yazd, Iran E-mail: htorabi@yazduni.ac.ir