Research Article Open Access

R. Pourmousa*, M. Rezapour, and M. Mashinchi

A theory for non-linear prediction approach in the presence of vague variables: with application to BMI monitoring

DOI 10.1515/demo-2015-0016
Received June 22, 2015; accepted November 4, 2015

Abstract: In the statistical literature, truncated distributions can be used for modeling real data. Due to error of measurement in truncated continuous data, choosing a crisp trimmed point caucuses a fault inference, so using fuzzy sets to define a threshold point may leads us more efficient results with respect to crisp thresholds. Arellano-Valle et al. [2] defined a selection distribution for analysis of truncated data with crisp threshold. In this paper, we define fuzzy multivariate selection distribution that is an extension of the selection distributions using fuzzy threshold. A practical data set with a fuzzy threshold point is considered to investigate the relationship between high blood pressure and BMI.

Keywords: Multivariate selection distribution, Membership function, Fuzzy event

MSC: 26E50, 60E05, 62J02

1 Introduction

Statistical analysis, in traditional form, is based on crispness of data, random variables, hypotheses, decision rules, parameters, and so on. But, information imprecision and uncertainty exist in real-world applications that can be caused by human errors in collecting data or some unexpected situations. Therefore, the fuzzy set theory naturally provides an appropriate tool in modeling the imprecise concepts. Suppose we should present a threshold that can be used in our decision. But, some vagueness in our data such as error of measurements may be distort our modeling. Therefore, using a fuzzy threshold can improve our decision. This leads to the idea of studying selection distribution under fuzzy events, which is an extension of conditional fuzzy probability introduced by Zadeh [19]. Pourmousa and Mashinchi [15] introduced a fuzzy method for producing family of univariate and multivariate skew-elliptical distributions based on fuzzy conditional events. They, also used the idea of fuzzy events for calculating tail conditional expectations for elliptical and skew-elliptical distributions. In fact, they considered the random vector $\mathbf{X}$ given the fuzzy event in which $\mathbf{X}$ is greater than a threshold.

In practical situation, demographic, behavioral, and physiological variables may be dependent if one of them is very low or high. For example, alcohol drinking and high blood pressure may have significant relationship that was studied in [17]. Association between blood pressure and body mass index in lean populations was discovered in [13]. Sodium intake among people with normal and high blood pressure was studied in Ajani [1]. The effects of iron supplementation on serum copper and zinc levels in pregnant women with high-

***Corresponding Author: R. Pourmousa:** Department of Statistics, Faculty of Mathematics and Computer, Shahid Bahonar University of Kerman, Kerman, Iran, E-mail: pourm@uk.ac.ir
M. Rezapour, M. Mashinchi: Department of Statistics, Faculty of Mathematics and Computer, Shahid Bahonar University of Kerman, Kerman, Iran, E-mail: mohsenrzp@uk.ac.ir, mashinchi@uk.ac.ir

normal hemoglobin was investigated in [20]. Association between parity and breastfeeding with maternal high blood pressure was discovered in [14].

Now suppose $(\mathbf{X}, \mathbf{Y})$ be a random vector and we are eager to predict $\mathbf{X}$, given that the random vector $\mathbf{Y}$ is greater (less) than a threshold, since a variable may be related with another one, if it be either less or greater than a threshold. Selection distributions can be used to consider these kinds of relationship in the classical statistics [2]. An important class of distributions in this case is selection elliptical distribution that can be used when data has elliptical distribution. Because of the vagueness in determination of crisp threshold, using fuzzy threshold motivates us to introduce fuzzy selection distribution as an extension of selection distribution.

The organization of this paper is as follows. In Section 2, we introduced multivariate selection distribution and then we present selection elliptical distribution considered in [2]. Unified skew-elliptical distribution function is considered as a special case of selection elliptical distributions. Membership function and probability of fuzzy events are also introduced in this section. In Section 3, using distribution of a random vector given a fuzzy event, we introduce fuzzy selection elliptical distribution. Unified multivariate fuzzy skew-elliptical distribution is also considered as a special case. In Section 4, we consider the non-linear regression given a fuzzy event. Lastly, a practical example on the relation between high blood pressure and Body Mass Index (BMI) is given, which applied the result reported in this paper.

2 Preliminaries

In this section, we briefly review the class of multivariate selection distributions introduced by Arellano-Valle et al. [2]. This class can be used for considering a random variable given another one belongs to a known set. For example, we can use it for considering a random variable given another one which is either less or greater a thresholds. Also, we give a brief introduction on the probability measure of fuzzy events introduced by L. A. Zadeh [18, 19].

2.1 Multivariate selection distributions

Let $\mathbf{X} \in \mathbb{R}^m$ and $\mathbf{Y} \in \mathbb{R}^n$ be two random vectors and C be a measurable subset of $\mathbb{R}^m$. Then, Arellano-Valle et al. [2] defined a selection distribution as the conditional distribution of $\mathbf{Y}$, given $\mathbf{X} \in C$. Specifically, a n-dimensional random vector $\mathbf{U}$ is said to have a multivariate selection distribution, denoted by $\mathbf{U} \sim \mathbf{SLCT}_{n,m}$, with parameters depending on the characteristics of $\mathbf{X}$, $\mathbf{Y}$ and C, if

$$\mathbf{U} \stackrel{d}{=} \mathbf{Y} \mid (\mathbf{X} \in C)\,. \tag{1}$$

These authors also showed that if the selection probability density function of $\mathbf{U}$ exists, then it is given by

$$f_{\mathbf{U}}(\mathbf{u}) = f_{\mathbf{Y}}(\mathbf{u})\frac{Pr(\mathbf{X} \in C|\mathbf{Y} = \mathbf{u})}{Pr(\mathbf{X} \in C)}, \quad \mathbf{u} \in \mathbb{R}^n\,. \tag{2}$$

One of the important cases of selection distributions is when $\mathbf{X}$ and $\mathbf{Y}$ are jointly elliptical distributed as

$$\begin{pmatrix} \mathbf{X} \\ \mathbf{Y} \end{pmatrix} \sim \mathbf{EC}_{m+n}\left(\mu = \begin{pmatrix} \mu_{\mathbf{X}} \\ \mu_{\mathbf{Y}} \end{pmatrix}, \Sigma = \begin{pmatrix} \Sigma_{\mathbf{XX}} & \Sigma_{\mathbf{YX}}^T \\ & \Sigma_{\mathbf{YY}} \end{pmatrix}, h^{(m+n)}\right), \tag{3}$$

where $\mu_{\mathbf{X}} \in \mathbb{R}^m$, $\mu_{\mathbf{Y}} \in \mathbb{R}^n$, $\Sigma_{\mathbf{YY}} \in \mathbb{R}^{n\times n}$, $\Sigma_{\mathbf{XX}} \in \mathbb{R}^{m\times m}$, $\Sigma_{\mathbf{YX}} \in \mathbb{R}^{n\times m}$ and $h^{(m+n)}$ is the density generator function. In this case, $\mathbf{U} \stackrel{d}{=} \mathbf{Y} \mid (\mathbf{X} \in C)$ is said to have a multivariate selection elliptical distribution, denoted by $\mathbf{U}_{\boldsymbol{\theta}} \sim \mathbf{SLCT} - \mathbf{EC}_{n,m}(\boldsymbol{\theta})$, where $\boldsymbol{\theta} = (\mu_{\mathbf{Y}}, \mu_{\mathbf{X}}, \Sigma_{\mathbf{YY}}, \Sigma_{\mathbf{XX}}, \Sigma_{\mathbf{YX}}, C)$.

In the special case when $C = \{\mathbf{x} \in \mathbb{R}^m \mid \mathbf{x} > \mathbf{0}\}$, then the probability density function (pdf) in (5) reduces to the pdf of the unified multivariate skew-elliptical distribution presented in [2], denoted by $\mathbf{U}_{\boldsymbol{\theta}} \sim \mathbf{SUE}_{n,m}(\boldsymbol{\theta})$, $\boldsymbol{\theta} = (\mu_{\mathbf{Y}}, \mu_{\mathbf{X}}, \Sigma_{\mathbf{YY}}, \Sigma_{\mathbf{XX}}, \Sigma_{\mathbf{YX}})$.

One of the important cases of selection elliptical distributions is the multivariate selection normal distribution, denoted by $\mathbf{U}_{\boldsymbol{\theta}} \sim \mathbf{SLCT} - \mathbf{N}_{n,m}(\boldsymbol{\theta})$, when $\mathbf{X}$ and $\mathbf{Y}$ are jointly normal distributed as

$$\begin{pmatrix} \mathbf{X} \\ \mathbf{Y} \end{pmatrix} \sim \mathbf{N}_{m+n}\left(\begin{pmatrix} \mu_{\mathbf{X}} \\ \mu_{\mathbf{Y}} \end{pmatrix}, \begin{pmatrix} \Sigma_{\mathbf{XX}} & \Sigma_{\mathbf{YX}}^T \\ & \Sigma_{\mathbf{YY}} \end{pmatrix}\right), \tag{4}$$

The pdf of $\mathbf{U}_{\boldsymbol{\theta}}$ can be derived as (see [3] for further details)

$$\phi_{\mathbf{SLCT}-\mathbf{N}_{n,m}}(\mathbf{u};\boldsymbol{\theta}) = \frac{\phi_n(\mathbf{u};\mu_{\mathbf{Y}}, \Sigma_{\mathbf{YY}})\, \Phi_m\left(C; \mu_{\mathbf{X}} + \Sigma_{\mathbf{YX}}^T \Sigma_{\mathbf{YY}}^{-1}(\mathbf{u}-\mu_{\mathbf{Y}}), \Sigma_{\mathbf{XX.Y}}\right)}{\Phi_m(C; \mu_{\mathbf{X}}, \Sigma_{\mathbf{XX}})}, \quad \mathbf{u} \in \mathbb{R}^n, \tag{5}$$

where $\Sigma_{\mathbf{XX.Y}} = \Sigma_{\mathbf{XX}} - \Sigma_{\mathbf{YX}}^T \Sigma_{\mathbf{YY}}^{-1} \Sigma_{\mathbf{YX}}$, $\phi_n(\cdot; \mu_{\mathbf{Y}}, \Sigma_{\mathbf{YY}})$ denotes the pdf of $\mathbf{N}_n(\mu_{\mathbf{Y}}, \Sigma_{\mathbf{YY}})$ and $\Phi_m(C; \mu_{\mathbf{X}}, \Sigma_{\mathbf{XX}})$ denotes $\Pr(\mathbf{X} \in C)$, where $X \sim \mathbf{N}_m(\mu_{\mathbf{X}}, \Sigma_{\mathbf{XX}})$. In the special case when $C = \{\mathbf{x} \in \mathbb{R}^m \mid \mathbf{x} > \mathbf{0}\}$, then the pdf in (5) reduces

to the pdf of the unified multivariate skew-normal distribution presented in [3], denoted by $\mathbf{U}_{\boldsymbol{\theta}} \sim \mathbf{SUN}_{n,m}(\boldsymbol{\theta})$, $\boldsymbol{\theta} = (\mu_{\mathbf{Y}}, \mu_{\mathbf{X}}, \Sigma_{\mathbf{YY}}, \Sigma_{\mathbf{XX}}, \Sigma_{\mathbf{YX}})$, given by

$$\phi_{SUN_{n,m}}(\mathbf{u};\boldsymbol{\theta}) = \frac{\phi_n(\mathbf{u}; \mu_{\mathbf{Y}}, \Sigma_{\mathbf{YY}})\, \Phi_m\left(\mu_{\mathbf{X}} + \Sigma_{\mathbf{YX}}^T \Sigma_{\mathbf{YY}}^{-1}(\mathbf{u} - \mu_{\mathbf{Y}}); \Sigma_{\mathbf{XX.Y}}\right)}{\Phi_m(\mu_{\mathbf{X}}; \Sigma_{\mathbf{XX}})}, \quad \mathbf{u} \in \mathbb{R}^n \ . \tag{6}$$

Its moment generating function (mgf) is

$$M_{\mathbf{SUN}_{n,m}}(\mathbf{t}; \theta) = \frac{\exp\left(\mu_{\mathbf{Y}}^T \mathbf{t} + \frac{1}{2}\mathbf{t}^T \Sigma_{\mathbf{YY}} \mathbf{t}\right) \Phi_m\left(\mu_{\mathbf{X}} + \Sigma_{\mathbf{YX}}^T \mathbf{t};\ \Sigma_{\mathbf{XX}}\right)}{\Phi_m(\mu_{\mathbf{X}}; \Sigma_{\mathbf{XX}})}, \quad \mathbf{t} \in \mathbb{R}^n, \tag{7}$$

where $\Phi_m(\cdot; \Sigma_{\mathbf{XX.Y}})$ and $\Phi_m(\cdot; \Sigma_{\mathbf{XX}})$ denote the cumulative distribution functions (cdf) of $\mathbf{N}_m(\mathbf{0}, \Sigma_{\mathbf{XX.Y}})$ and $\mathbf{N}_m(\mathbf{0}; \Sigma_{\mathbf{XX}})$, respectively.

2.2 Concepts on fuzzy sets

The concept of fuzzy set was initiated by Zadeh [18] in 1965. Let Ω be a sample space of a random experiment and $\mathcal{A}$ a fuzzy subset of Ω. If for all $x \in \Omega$, there is a number $\mu_{\mathcal{A}}(x) \in [0, 1]$ assigned to represent the membership of x to $\mathcal{A}$, then $\mu_{\mathcal{A}}$ is called the membership function of $\mathcal{A}$. A good overview of the various interpretations of membership functions in fuzzy set theory can be found in Dubois and Prade [10].

Let $(\Omega, \mathcal{F}, P)$ be a probability space and $\mathbf{X} \in \mathbb{R}^n$ be a random vector in Ω. A fuzzy event is a fuzzy set $\mathcal{A}$ in Ω whose membership function, $\mu_{\mathcal{A}} : \Omega \longrightarrow [0, 1]$, is Borel measurable function and probability of the fuzzy event is defined by

$$\Pr(\mathcal{A}) = \int_{\Omega} \mu_{\mathcal{A}}(\mathbf{x}) dP(\mathbf{x}). \tag{8}$$

Based on the formula for the probability of a fuzzy set in Eq. (8), Zadeh [19] defined the conditional distribution of $\mathbf{X}$ given the fuzzy event $\mathcal{A}$, as

$$f_{\mathbf{X}|\mathcal{A}}(\mathbf{x}) = \frac{\mu_{\mathcal{A}}(\mathbf{x}) dP(\mathbf{x})}{\int_{\mathbf{x}\in\Omega} \mu_{\mathcal{A}}(\mathbf{x}) dP(\mathbf{x})}, \quad \mathbf{x} \in \Omega\,. \tag{9}$$

3 Fuzzy selection distributions

Our purpose in this paper is to define a parallel notion to the class of multivariate selection distributions, based on fuzzy conditional events. This can be used in many practical applications. For example, consider the relationship between BMI and blood pressure. In the literature, researcher has been considered the relation between either low or high BMI and blood pressure [11]. Because of the vagueness implied by the exact determination of these thresholds, considering these events as a crisp phenomenon lead some misleading inferences. We introduce fuzzy selection distributions as an extension of selection distributions to consider an event given another fuzzy event. Let $(\mathbf{X}, \mathbf{Y})$ be a random vector and $\mathcal{C}$ a measurable fuzzy set $\sigma(\mathbf{Y})$, then we want to consider the distribution of $\mathbf{X}$ given $\mathbf{Y} \in \mathcal{C}$ which is named multivariate fuzzy selection distribution (FSLC) throughout this paper.

Definition 3.1. (Multivariate Fuzzy Selection Distribution). Let $\mathbf{X} \in \mathbb{R}^m$ and $\mathbf{Y} \in \mathbb{R}^n$ be two random vectors and $\mathcal{C}$ be a fuzzy subset of $\mathbb{R}^m$ with the membership function $\mu_{\mathcal{C}}$. Then, the random vector $\mathbf{V}$ is said to have a multivariate fuzzy selection distribution, denoted by $\mathbf{V} \sim \mathbf{FSLCT}_{n,m}$, with parameters depending on the characteristics of $\mathbf{X}$, $\mathbf{Y}$ and $\mathcal{C}$, if $\mathbf{V} \overset{d}{=} \mathbf{Y} \mid (\mathbf{X} \in \mathcal{C})$.

In the following result, we obtain pdf of the fuzzy selection random vector $\mathbf{V}$. Note that this is an extension of their corresponding result due to the selection distributions in [3], i.e. when $\mu_{\mathcal{C}}(\mathbf{x})$ be an indicator function, the pdf of the selection distributions follows.

Theorem 3.2. *Let $\mathbf{X} \in \mathbb{R}^m$ and $\mathbf{Y} \in \mathbb{R}^n$ be two random vectors with joint pdf $f_{\mathbf{X},\mathbf{Y}}$ with respect to probability measure (pm) v and marginals $f_{\mathbf{X}}$ with respect to pm v_1 and $f_{\mathbf{Y}}$ with respect to pm v_2. The pdf of $\mathbf{V} \overset{d}{=} \mathbf{Y} \mid (\mathbf{X} \in \mathcal{C})$ is*

$$f_{\mathbf{FSLCT}_{n,m}}(\mathbf{y}) = \frac{1}{\Pr(\mathbf{X} \in \mathcal{C})} f_{\mathbf{Y}}(\mathbf{y}) E\left(\mu_{\mathcal{C}}(\mathbf{X}) \,|\, \mathbf{Y} = \mathbf{y}\right), \quad \mathbf{y} \in \mathbb{R}^n \tag{10}$$

where $Pr(\mathbf{X} \in \mathcal{C}) = E(\mu_{\mathcal{C}}(\mathbf{X})) = \int_{\mathbf{x}\in\mathbb{R}^m} f_{\mathbf{X}}(\mathbf{x})\, \mu_{\mathcal{C}}(\mathbf{x})\, d\mathbf{x}$ and

$$E\left(\mu_{\mathcal{C}}(\mathbf{X}) \,|\, \mathbf{Y} = \mathbf{y}\right) = \int_{\mathbf{x}\in\mathbb{R}^m} f_{\mathbf{X}|\mathbf{Y}=\mathbf{y}}(\mathbf{x})\, \mu_{\mathcal{C}}(\mathbf{x})\, dv_1(\mathbf{x}),$$

is the conditional expectation of the random vector $\mu_{\mathcal{C}}(\mathbf{X})$ given $(\mathbf{Y} = \mathbf{y})$, and $\mu_C(\cdot)$ defined in Definition (3.1).

Proof. Suppose that $F_{\mathbf{FSLCT}_{n,m}}$ and $F_{\mathbf{Y}}$ are cdfs of $\mathbf{V}$ and $\mathbf{Y}$, respectively. Then, we have

$$F_{\mathbf{FSLCT}_{n,m}}(\mathbf{y}) = Pr\left(\mathbf{Y} \le \mathbf{y} \,|\, \mathbf{X} \in \mathcal{C}\right) = \frac{Pr\left(\mathbf{Y} \le \mathbf{y}, \mathbf{X} \in \mathcal{C}\right)}{Pr\left(\mathbf{X} \in \mathcal{C}\right)} = F_{\mathbf{Y}}(\mathbf{y}) \frac{Pr\left(\mathbf{X} \in \mathcal{C} \,|\, \mathbf{Y} \le \mathbf{y}\right)}{Pr\left(\mathbf{X} \in \mathcal{C}\right)}.$$

But, we know that for $\mathbf{A} \in \mathbb{R}^n$, we have $P(\mathbf{X} \in \mathbf{A}|\mathbf{Y} \le \mathbf{y}) = \int_{\mathbf{A}} f_{\mathbf{X}|(\mathbf{Y}\le\mathbf{y})}(\mathbf{x})\, dv_1(\mathbf{x})$ with

$$f_{\mathbf{X}|(\mathbf{Y}\le\mathbf{y})}(\mathbf{x}) = f_{\mathbf{X}}(\mathbf{x}) \frac{F_{\mathbf{Y}|\mathbf{X}}(\mathbf{y}\,|\,\mathbf{x})}{F_{\mathbf{Y}}(\mathbf{y})}, \quad \mathbf{x} \in \mathbb{R}^m, \tag{11}$$

where $F_{\mathbf{Y}|\mathbf{X}}$ is cdf of $\mathbf{Y}$ given $\mathbf{X} = \mathbf{x}$. By using Eq. (8) and Theorem 16.11 in [9], we obtain

$$Pr(\mathbf{X} \in \mathcal{C}|\mathbf{Y} \le \mathbf{y}) = \int_{\mathbf{x}\in\mathbb{R}^m} f_{\mathbf{X}|(\mathbf{Y}\le\mathbf{y})}(\mathbf{x})\, \mu_{\mathcal{C}}(\mathbf{x})\, dv_1(\mathbf{x}).$$

Therefore, we have

$$F_{\mathbf{FSLCT}_{n,m}}(\mathbf{y}) = \frac{F_{\mathbf{Y}}(\mathbf{y})}{Pr\left(\mathbf{X} \in \mathcal{C}\right)} \int_{\mathbf{x}\in\mathbb{R}^m} f_{\mathbf{X}|(\mathbf{Y}\le\mathbf{y})}(\mathbf{x})\, \mu_{\mathcal{C}}(\mathbf{x})\, dv_1(\mathbf{x}), \tag{12}$$

By taking into consideration (11) in (12) and differentiating with respect to $\mathbf{y}$, the proof is completed. □

Remark 3.3. Note that the pdf of the fuzzy selection random vector $\mathbf{V}$ given in (9) and (10) reduces to (2) when $\mu_{\mathcal{C}}(\mathbf{x})$ be an indicator function.

Consider the special case in which $\mathbf{X}$ and $\mathbf{Y}$ have joint elliptical distribution as in Eq. (3). Then, the random vector $\mathbf{V} \stackrel{d}{=} (\mathbf{Y} \mid \mathbf{X} \in \mathcal{C})$ is said to have a multivariate fuzzy selection elliptical distribution which is denoted by $\mathbf{V} \sim \mathbf{FSLCT-EC}_{n,m}(\boldsymbol{\theta})$, where $\boldsymbol{\theta} = (\mu_{\mathbf{Y}}, \mu_{\mathbf{X}}, \Sigma_{\mathbf{YY}}, \Sigma_{\mathbf{XX}}, \Sigma_{\mathbf{YX}}, \mathcal{C})$. When $\mathcal{C} = \{\mathbf{x} \in \mathbb{R}^m \mid \mathbf{x} \gtrsim \mathbf{0}\}$, where "$\gtrsim$" states "almost equal bigger", then the pdf of $\mathbf{V}$ reduces to the pdf of the unified multivariate fuzzy skew-elliptical distribution which is denoted by $\mathbf{V} \sim \mathbf{FSUE}_{n,m}(\boldsymbol{\theta})$ where $\boldsymbol{\theta} = (\mu_{\mathbf{Y}}, \mu_{\mathbf{X}}, \Sigma_{\mathbf{YY}}, \Sigma_{\mathbf{XX}}, \Sigma_{\mathbf{YX}})$.

If $\mathbf{X}$ and $\mathbf{Y}$ have joint normal distribution as in Eq.(4), then $\mathbf{V} \stackrel{d}{=} (\mathbf{Y} \mid \mathbf{X} \in \mathcal{C})$ is said to have a multivariate fuzzy selection normal distribution which is denoted by $\mathbf{V} \sim \mathbf{FSLCT-N}_{n,m}(\boldsymbol{\theta})$. By Theorem 3.2, we can derive the pdf of $\mathbf{V}$ as

$$\phi_{\mathbf{FSLCT-N}_{n,m}}(\mathbf{y};\boldsymbol{\theta}) = \frac{1}{\Pr(\mathbf{X} \in \mathcal{C})} \phi_n\left(\mathbf{y}; \mu_{\mathbf{Y}}, \Sigma_{\mathbf{YY}}\right) \int_{\mathbf{x} \in \mathbb{R}^m} \phi_m\left(\mathbf{x}; \mu_{\mathbf{X}} + \Sigma_{\mathbf{YX}}^T \Sigma_{\mathbf{YY}}^{-1}(\mathbf{y} - \mu_{\mathbf{Y}}), \Sigma_{\mathbf{XX.Y}}\right) \mu_{\mathcal{C}}(\mathbf{x})\, d\mathbf{x}, \tag{13}$$

where $\Sigma_{\mathbf{XX.Y}} = \Sigma_{\mathbf{XX}} - \Sigma_{\mathbf{YX}}^T \Sigma_{\mathbf{YY}}^{-1} \Sigma_{\mathbf{YX}}$ and $\Pr(\mathbf{X} \in \mathcal{C}) = E(\mu_{\mathcal{C}}(\mathbf{X}))$ with $X \sim N_m(\mu_{\mathbf{X}}, \Sigma_{\mathbf{XX}})$. Now, consider a special case in which $m = n = 1$ and it will be used in the next section. For this purpose, let

$$\begin{pmatrix} X \\ Y \end{pmatrix} \sim \mathbf{N}_2\left(\begin{pmatrix} \mu_X \\ \mu_Y \end{pmatrix}, \begin{pmatrix} \sigma_X^2 & \sigma_{X.Y} \\ & \sigma_Y^2 \end{pmatrix}\right), \tag{14}$$

therefore, the pdf (13) reduces to

$$\phi_{FSLCT-N}(y;\theta) = \frac{1}{\Pr(X \in \mathcal{C})} \phi\left(y; \mu_Y, \sigma_Y^2\right) \int_{x \in \mathbb{R}} \phi\left(x; \alpha y + \beta, \sigma_{X.Y}^2\right) \mu_{\mathcal{C}}(x)\, dx, \tag{15}$$

where $\sigma_{X.Y}^2 = \sigma_X^2\left(1 - \rho^2\right)$, $\rho = \frac{\sigma_{X.Y}}{\sigma_X \sigma_Y}$, $\alpha = \rho \frac{\sigma_X}{\sigma_Y}$,, $\beta = \mu_X - \alpha \mu_Y$, and $\Pr(X \in \mathcal{C}) = E(\mu_{\mathcal{C}}(X))$ with $X \sim N(\mu_X, \sigma_X^2)$.

Note that if the membership function is an indicator function as follow

$$\mu_{\mathcal{C}}(x) = \begin{cases} 0, & x < 0; \\ 1, & x \geq 0. \end{cases} \tag{16}$$

then, we have the so-called extended skew-normal pdf given by

$$\phi_{FSLCT-N}(y;\theta) = \frac{1}{\Phi(\frac{\mu}{\sigma_X})} \phi\left(\frac{y - \mu_Y}{\sigma_{YY}}\right) \Phi\left(\frac{\alpha y + \beta}{\sigma_{X.Y}}\right), \tag{17}$$

where ϕ and Φ are pdf and cdf of standard normal distribution, respectively. Now, if $\mu_X = \mu_Y = 0$ and $\sigma_X = \sigma_Y = 1$, (17) reduces to the well-known skew-normal pdf given by

$$\phi_{FSLCT-N}(y;\theta) = 2\phi(y)\,\Phi(\lambda y); \quad \lambda = \frac{\rho}{\sqrt{1 - \rho^2}}.$$

4 Prediction via fuzzy selection distributions

In practical situation, we want to know the mean value of a random variable given another variable belongs to a certain set. For example, the mean value of blood pressure for people with low or high BMI are desirable, if we use fuzzy events we may obtain a better conclusion for determination of threshold which cause a more reliable prediction of dependent variables. In this section, we obtain the mathematical expectation of random variable $V_a \stackrel{d}{=} Y \mid (X \in \mathcal{C}_a)$ where the random vector $(X, Y)^T$ has joint distribution function (14) and fuzzy event $\mathcal{C}_a = \{x \in \mathbb{R} \mid x \succcurlyeq a\}$ with the following membership function:

$$\mu_{\mathcal{C}_a}(x) = \begin{cases} 0, & x < a - 1; \\ x - a + 1, & a - 1 \leq x < a; \\ 1, & x \geq a. \end{cases} \tag{18}$$

By (15) the density function of the random variable V_a equals

$$\phi_{FSLCT-N}(y;\theta) = \phi\left(y;\mu_Y,\sigma_Y^2\right)\frac{E\left(\mu_{\mathcal{C}_a}(X)|Y=y\right)}{E\left(\mu_{\mathcal{C}_a}(X)\right)}, \tag{19}$$

where $X|Y = y \sim N(\mu_{X\cdot Y}, \sigma^2_{X\cdot Y})$ and $X \sim N(\mu_X, \sigma^2_X)$. Using the following equality

$$\int_{\xi_1}^{\xi_2}(x+\lambda)\,\phi\left(x;\mu,\sigma^2\right)dx = (\mu+\lambda)\left(\Phi\left(\frac{\xi_2-\mu}{\sigma}\right)-\Phi\left(\frac{\xi_1-\mu}{\sigma}\right)\right)-\sigma\left(\phi\left(\frac{\xi_2-\mu}{\sigma}\right)-\phi\left(\frac{\xi_1-\mu}{\sigma}\right)\right),$$

for each $\lambda, \mu \in \mathbb{R}$, $\sigma \in \mathbb{R}^+$, and $\xi_1 < \xi_2 \in \mathbb{R}$ in (19), the density function of the random variable V_a reduces to

$$\begin{aligned}\phi_{FSLCT-N}(y;\theta) = \frac{\phi\left(\frac{y-\mu_Y}{\sigma_Y}\right)}{\sigma_Y \Pr(X\in\mathcal{C}_a)}&\left(\sigma_{X.Y}\left(\phi\Big(\frac{a-1-\alpha y-\beta}{\sigma_{X.Y}}\Big)-\phi\Big(\frac{a-\alpha y-\beta}{\sigma_{X.Y}}\Big)\right)\right.\\ &\left.+(\alpha y+\beta+1-a)\left(\Phi\Big(\frac{a-\alpha y-\beta}{\sigma_{X.Y}}\Big)-\Phi\Big(\frac{a-1-\alpha y-\beta}{\sigma_{X.Y}}\Big)\right)+\Phi\Big(\frac{\alpha y+\beta-a}{\sigma_{X.Y}}\Big)\right),\end{aligned} \tag{20}$$

where α, β, and $\sigma_{X.Y}$ are presented in (15) and

$$\Pr(X\in\mathcal{C}_a) = (\mu_X-a+1)\left(\Phi(\frac{a-\mu_X}{\sigma_X})-\Phi(\frac{a-1-\mu_X}{\sigma_X})\right)-\sigma_X\left(\phi(\frac{a-\mu_X}{\sigma_X})-\phi(\frac{a-1-\mu_X}{\sigma_X})\right)+\Phi(\frac{\mu_X-a}{\sigma_X}). \tag{21}$$

The last term is obtained by the equalities (8) and (20). The following lemma that may be obtain with a straightforward calculation can be used to compute the mathematical expectation of the random variable V_a with pdf (20).

Lemma 4.1. *For any $A, C \in \mathbb{R}^+$ and $B, D \in \mathbb{R}$ we have*

$$\int_{-\infty}^{+\infty}\phi(Ax-B)\,\Phi(Cx-D)\,dx = \frac{1}{A}\Phi\left(\frac{BC-AD}{\sqrt{A^2+C^2}}\right),$$

$$\int_{-\infty}^{+\infty}x\phi(Ax-B)\,\phi(Cx-D)\,dx = \frac{AB+CD}{\left(A^2+C^2\right)^{\frac{3}{2}}}\phi\left(\frac{BC-AD}{\sqrt{A^2+C^2}}\right),$$

$$\int_{-\infty}^{+\infty}x\phi(Ax-B)\,\Phi(Cx-D)\,dx = \frac{B}{A^2}\Phi\left(\frac{BC-AD}{\sqrt{A^2+C^2}}\right)+\frac{C}{A^2\sqrt{A^2+C^2}}\phi\left(\frac{BC-AD}{\sqrt{A^2+C^2}}\right),$$

$$\int_{-\infty}^{+\infty}x^2\phi(Ax-B)\,\Phi(Cx-D)\,dx = \frac{1+B^2}{A^3}\Phi\left(\frac{BC-AD}{\sqrt{A^2+C^2}}\right)+\frac{C\left(2A^2B+BC^2+ACD\right)}{A^3\left(A^2+C^2\right)^{\frac{3}{2}}}\phi\left(\frac{BC-AD}{\sqrt{A^2+C^2}}\right).$$

In the following theorem, we present a compact form for mathematical expectation of V_a with pdf (19) that is

$$E(V_a) = \frac{E\left(Y\left(\mu_{\mathcal{C}_a}(X)|Y=y\right)\right)}{E\left(\mu_{\mathcal{C}_a}(X)\right)}, \tag{22}$$

where $X|Y = y$ and X distributed as in (19) and $Y \sim N(\mu_Y, \sigma_Y^2)$. We will use this compact form in further applications.

Theorem 4.2. *The mathematical expectation of the fuzzy selection random variable V_a with pdf (19) is*

$$\begin{aligned}E(V_a) = \frac{1}{\Pr(X\in\mathcal{C}_a)}&\left(\sigma_{X.Y}\Big(\gamma_1(a-1)-\gamma_1(a)\Big)+\alpha\Big(\gamma_2(a-1)-\gamma_2(a)\Big)\right.\\ &\left.+\Big(\beta+1-a\Big)\Big(\gamma_3(a-1)-\gamma_3(a)\Big)+\mu_Y-\gamma_3(a)\right),\end{aligned} \tag{23}$$

where $\Pr(X \in \mathcal{C}_a)$ *is given in* (21) *and*

$$\begin{aligned}
\gamma_1(a) &= \frac{\sigma^3_{X.Y}}{\sigma^3_X}\left(\mu_Y - \frac{\sigma_{XY}(\beta - a)}{\sigma^2_{X.Y}}\right)\phi\left(\frac{a-\mu_X}{\sigma_X}\right), \\
\gamma_2(a) &= \left(\mu_Y^2 + \sigma_Y^2\right)\Phi\left(-\frac{a-\mu_X}{\sigma_X}\right) - \frac{2\alpha}{\sigma_{X.Y}}\left(\mu_Y - \frac{\sigma_Y}{2\sigma^2_{X.Y}}\alpha(\beta - a - \alpha\mu_Y)\right)\phi\left(\frac{a-\mu_X}{\sigma_X}\right), \\
\gamma_3(a) &= \mu_Y\Phi\left(-\frac{a-\mu_X}{\sigma_X}\right) - \rho\sigma_Y\,\phi\left(\frac{a-\mu_X}{\sigma_X}\right).
\end{aligned}$$

Proof. We can obtain the desirable expectation as $E(V_a) = \int_{x\in\mathbb{R}} y\phi_{FSLCT-N}(y)dy$, where $\phi_{FSLCT-N}(y)$ is computed in (20). Therefore, we have

$$\begin{aligned}
E(V_a) = \frac{1}{Pr(X \in \mathcal{C}_a)}\Bigg[&\int_{-\infty}^{+\infty}\frac{y}{\sigma_Y}\phi\left(\frac{y-\mu_Y}{\sigma_Y}\right)\left(\sigma_{X.Y}\left(\phi\left(\frac{a-1-\alpha y-\beta}{\sigma_{X.Y}}\right) - \phi\left(\frac{a-\alpha y-\beta}{\sigma_{X.Y}}\right)\right)\right)dy \\
&+ (\alpha y + \beta - a + 1)\int_{-\infty}^{+\infty}\frac{y}{\sigma_Y}\phi\left(\frac{y-\mu_Y}{\sigma_Y}\right)\left(\Phi\left(\frac{a-\alpha y-\beta}{\sigma_{X.Y}}\right) - \Phi\left(\frac{a-1-\alpha y-\beta}{\sigma_{X.Y}}\right)\right)dy \\
&+ \int_{-\infty}^{+\infty}\frac{y}{\sigma_Y}\phi\left(\frac{y-\mu_Y}{\sigma_Y}\right)\Phi\left(\frac{\alpha y+\beta-a}{\sigma_{X.Y}}\right)dy\Bigg]. \qquad (24)
\end{aligned}$$

By Lemma 4.1, we can rewrite (24) as

$$\begin{aligned}
&\frac{\sigma_{X.Y}}{\sigma_Y Pr(X\in\mathcal{C}_a)}\left(\frac{\frac{1}{\sigma_Y}\frac{\mu_Y}{\sigma_Y} - \frac{\alpha}{\sigma_{X.Y}}\frac{a-1-\beta}{\sigma_{X.Y}}}{\left((\frac{1}{\sigma_Y})^2 + (\frac{\alpha}{\sigma_{X.Y}})^2\right)^{\frac{3}{2}}}\phi\left(-\frac{\frac{\mu_Y}{\sigma_Y}\frac{\alpha}{\sigma_{X.Y}} - \frac{1}{\sigma_Y}\frac{a-1-\beta}{\sigma_{X.Y}}}{\sqrt{(\frac{1}{\sigma_Y})^2 + (\frac{\alpha}{\sigma_{X.Y}})^2}}\right)\right. \\
&\qquad\left. - \frac{\frac{1}{\sigma_Y}\frac{\mu_Y}{\sigma_Y} - \frac{\alpha}{\sigma_{X.Y}}\frac{a-\beta}{\sigma_{X.Y}}}{\left((\frac{1}{\sigma_Y})^2 + (\frac{\alpha}{\sigma_{X.Y}})^2\right)^{\frac{3}{2}}}\phi\left(-\frac{\frac{\mu_Y}{\sigma_Y}\frac{\alpha}{\sigma_{X.Y}} - \frac{1}{\sigma_Y}\frac{a-\beta}{\sigma_{X.Y}}}{\sqrt{(\frac{1}{\sigma_Y})^2 + (\frac{\alpha}{\sigma_{X.Y}})^2}}\right)\right) \\
&+\frac{\alpha}{\sigma_Y Pr(X\in\mathcal{C}_a)}\left(\frac{1+(\frac{\mu_Y}{\sigma_Y})^2}{(\frac{1}{\sigma_Y})^3}\Phi\left(\frac{-\frac{\mu_Y}{\sigma_Y}\frac{\alpha}{\sigma_{X.Y}} - \frac{1}{\sigma_Y}\frac{a-\beta}{\sigma_{X.Y}}}{\sqrt{(\frac{1}{\sigma_Y})^2 + (\frac{\alpha}{\sigma_{X.Y}})^2}}\right)\right. \\
&\qquad+\frac{-\frac{\alpha}{\sigma_{X.Y}}\left(2(\frac{1}{\sigma_Y})^2\frac{\mu_Y}{\sigma_Y} + \frac{\mu_Y}{\sigma_Y}(\frac{\alpha}{\sigma_{X.Y}})^2 + \frac{1}{\sigma_Y}\frac{\alpha}{\sigma_{X.Y}}\frac{a-\beta}{\sigma_{X.Y}}\right)}{(\frac{1}{\sigma_Y})^3\left((\frac{1}{\sigma_Y})^2 + (\frac{\alpha}{\sigma_{X.Y}})^2\right)^{\frac{3}{2}}}\phi\left(\frac{\frac{\mu_Y}{\sigma_Y}\frac{\alpha}{\sigma_{X.Y}} - \frac{1}{\sigma_Y}\frac{a-\beta}{\sigma_{X.Y}}}{\sqrt{(\frac{1}{\sigma_Y})^2 + (\frac{\alpha}{\sigma_{X.Y}})^2}}\right) \\
&\qquad-\frac{1+(\frac{\mu_Y}{\sigma_Y})^2}{(\frac{1}{\sigma_Y})^3}\Phi\left(\frac{-\frac{\mu_Y}{\sigma_Y}\frac{\alpha}{\sigma_{X.Y}} - \frac{1}{\sigma_Y}\frac{a-1-\beta}{\sigma_{X.Y}}}{\sqrt{(\frac{1}{\sigma_Y})^2 + (\frac{\alpha}{\sigma_{X.Y}})^2}}\right) \\
&\qquad\left.-\frac{-\frac{\alpha}{\sigma_{X.Y}}\left(2(\frac{1}{\sigma_Y})^2\frac{\mu_Y}{\sigma_Y} + \frac{\mu_Y}{\sigma_Y}(\frac{\alpha}{\sigma_{X.Y}})^2 + \frac{1}{\sigma_Y}\frac{\alpha}{\sigma_{X.Y}}\frac{a-1-\beta}{\sigma_{X.Y}}\right)}{(\frac{1}{\sigma_Y})^3\left((\frac{1}{\sigma_Y})^2 + (\frac{\alpha}{\sigma_{X.Y}})^2\right)^{\frac{3}{2}}}\phi\left(\frac{\frac{\mu_Y}{\sigma_Y}\frac{\alpha}{\sigma_{X.Y}} - \frac{1}{\sigma_Y}\frac{a-1-\beta}{\sigma_{X.Y}}}{\sqrt{(\frac{1}{\sigma_Y})^2 + (\frac{\alpha}{\sigma_{X.Y}})^2}}\right)\right) \\
&+\frac{\beta-a+1}{\sigma_Y Pr(X\in\mathcal{C}_a)}\left(\frac{\frac{\mu_Y}{\sigma_Y}}{(\frac{1}{\sigma_Y})^2}\Phi\left(\frac{-\frac{\mu_Y}{\sigma_Y}\frac{\alpha}{\sigma_{X.Y}} - \frac{1}{\sigma_Y}\frac{a-\beta}{\sigma_{X.Y}}}{\sqrt{(\frac{1}{\sigma_Y})^2 + (\frac{\alpha}{\sigma_{X.Y}})^2}}\right) + \frac{-\frac{\alpha}{\sigma_{X.Y}}}{(\frac{1}{\sigma_Y})^2\sqrt{(\frac{1}{\sigma_Y})^2 + (\frac{\alpha}{\sigma_{X.Y}})^2}}\phi\left(\frac{-\frac{\mu_Y}{\sigma_Y}\frac{\alpha}{\sigma_{X.Y}} - \frac{1}{\sigma_Y}\frac{a-\beta}{\sigma_{X.Y}}}{\sqrt{(\frac{1}{\sigma_Y})^2 + (\frac{\alpha}{\sigma_{X.Y}})^2}}\right)\right. \\
&\qquad\left.-\frac{\frac{\mu_Y}{\sigma_Y}}{(\frac{1}{\sigma_Y})^2}\Phi\left(\frac{-\frac{\mu_Y}{\sigma_Y}\frac{\alpha}{\sigma_{X.Y}} - \frac{1}{\sigma_Y}\frac{a-1-\beta}{\sigma_{X.Y}}}{\sqrt{(\frac{1}{\sigma_Y})^2 + (\frac{\alpha}{\sigma_{X.Y}})^2}}\right) - \frac{-\frac{\alpha}{\sigma_{X.Y}}}{(\frac{1}{\sigma_Y})^2\sqrt{(\frac{1}{\sigma_Y})^2 + (\frac{\alpha}{\sigma_{X.Y}})^2}}\phi\left(\frac{-\frac{\mu_Y}{\sigma_Y}\frac{\alpha}{\sigma_{X.Y}} - \frac{1}{\sigma_Y}\frac{a-1-\beta}{\sigma_{X.Y}}}{\sqrt{(\frac{1}{\sigma_Y})^2 + (\frac{\alpha}{\sigma_{X.Y}})^2}}\right)\right) \\
&+\frac{1}{\sigma_Y Pr(X\in\mathcal{C}_a)}\left(\frac{\frac{\mu_Y}{\sigma_Y}}{(\frac{1}{\sigma_Y})^2}\Phi\left(\frac{\frac{\mu_Y}{\sigma_Y}\frac{\alpha}{\sigma_{X.Y}} + \frac{1}{\sigma_Y}\frac{a-\beta}{\sigma_{X.Y}}}{\sqrt{(\frac{1}{\sigma_Y})^2 + (\frac{\alpha}{\sigma_{X.Y}})^2}}\right) + \frac{\frac{\alpha}{\sigma_{X.Y}}}{(\frac{1}{\sigma_Y})^2\sqrt{(\frac{1}{\sigma_Y})^2 + (\frac{\alpha}{\sigma_{X.Y}})^2}}\phi\left(\frac{\frac{\mu_Y}{\sigma_Y}\frac{\alpha}{\sigma_{X.Y}} + \frac{1}{\sigma_Y}\frac{a-\beta}{\sigma_{X.Y}}}{\sqrt{(\frac{1}{\sigma_Y})^2 + (\frac{\alpha}{\sigma_{X.Y}})^2}}\right)\right).
\end{aligned}$$

By (15) and $(\frac{1}{\sigma_Y})^2 + (\frac{\alpha}{\sigma_{X.Y}})^2 = \frac{\sigma_X^2}{\sigma_Y^2\sigma_{X.Y}^2}$, we have

$$\frac{\frac{\mu_Y}{\sigma_Y}\frac{\alpha}{\sigma_{X.Y}} + \frac{1}{\sigma_Y}\frac{\sigma_Y^2\alpha(\beta-a)}{\sigma_{X.Y}^2}}{\sqrt{(\frac{1}{\sigma_Y})^2 + (\frac{\alpha}{\sigma_{X.Y}})^2}} = \frac{(\alpha\mu_Y + a - \beta)}{\sigma_X} = \frac{a - \mu_X}{\sigma_X} . \tag{25}$$

Similarly, we have

$$\frac{1}{\sigma_Y}\frac{\frac{1}{\sigma_Y}\frac{\mu_Y}{\sigma_Y} - \frac{\alpha}{\sigma_{X.Y}}\frac{a-\beta}{\sigma_{X.Y}}}{\left((\frac{1}{\sigma_Y})^2 + (\frac{\alpha}{\sigma_{X.Y}})^2\right)^{\frac{3}{2}}} = \frac{\sigma_{X.Y}^3}{\sigma_X^3}\left(\mu_Y - \frac{\sigma_{XY}}{\sigma_{X.Y}^2}(\beta - a)\right) , \tag{26}$$

$$\frac{-\frac{\alpha}{\sigma_{X.Y}}\left(2(\frac{1}{\sigma_Y})^2\frac{\mu_Y}{\sigma_Y} + \frac{\mu_Y}{\sigma_Y}(\frac{\alpha}{\sigma_{X.Y}})^2 + \frac{1}{\sigma_Y}\frac{\alpha}{\sigma_{X.Y}}\frac{a-\beta}{\sigma_{X.Y}}\right)}{\sigma_Y(\frac{1}{\sigma_Y})^3\left((\frac{1}{\sigma_Y})^2 + (\frac{\alpha}{\sigma_{X.Y}})^2\right)^{\frac{3}{2}}} = -\frac{2\alpha}{\sigma_{X.Y}}\left(\mu_Y - \frac{\sigma_Y}{2\sigma_{X.Y}^2}\alpha(\beta - a - \alpha\mu_Y)\right) , \tag{27}$$

and

$$\frac{\frac{\alpha}{\sigma_{X.Y}}}{\sigma_Y(\frac{1}{\sigma_Y})^2\sqrt{(\frac{1}{\sigma_Y})^2 + (\frac{\alpha}{\sigma_{X.Y}})^2}} = \rho\sigma_Y . \tag{28}$$

Therefore, by (25), (26), (27), and (28), the expectation of the random variable V_a reduces to (23) and this completes the proof of the theorem. □

4.1 Non-linear regression based on fuzzy selection distribution

In the simple regression model we assume that the response variable is a normal random variable, therefore the best linear prediction of mean value of response variable is its conditional expectation given dependent variable equal to a certain value. Now, consider the case in which $(\mathbf{X}, \mathbf{Y})$ be two random vector and we want to predict average value of $\mathbf{Y}$, given the random vector $\mathbf{X}$ equal to a certain value. For example, we want to predict mean yields of soybean plants (gms per plant) given the indicated levels of ozone over the growing season equal to certain value. Because of error in ozone measurement, it seems better to consider the given event as a fuzzy event. Here, instead of considering the given event as crisp event, we want to consider the fuzzy event in which independent variable is approximately equal to a certain value. Therefore, we should consider the mathematical expectation of dependent variable given the independent variable belong to a fuzzy set. In the classical statistical modeling, when two random vectors $(\mathbf{X}, \mathbf{Y})$ have jointly multivariate normal distributions then the conditional distribution function of $\mathbf{Y}$ given $\mathbf{X}$ belong to a Borel measurable set C is called selection normal distribution. In this subsection, we consider the non-linear regression based on fuzzy selection distribution. Suppose that the random vector (X, Y) has pdf (14). We can compute the non-linear regression based on fuzzy selection distribution by $E(Y|X \approx t)$. In fact, we should compute $E(Y|X \in \mathcal{C}_t)$, when $\mathcal{C}_t = \{x \in \mathbb{R} \,|\, x \approx t\}$, where $t \in \mathbb{R}$ is a fixed value and its membership function be

$$\mu_{\mathcal{C}_t}(x) = \begin{cases} 1 + x - t, & t - 1 \le x < t; \\ 1 + t - x, & t \le x < t + 1; \\ 0, & \text{otherwise}. \end{cases} \tag{29}$$

Therefore, the pdf of $V_t = Y|(X \in \mathcal{C}_t)$ equals

$$\phi_{FSLCT-N}(y) = \phi\left(y; \mu_Y, \sigma_Y^2\right)\frac{E\left(\mu_{\mathcal{C}_t}(X)|Y = y\right)}{E\left(\mu_{\mathcal{C}_t}(X)\right)} , \tag{30}$$

where $X|Y = y$ and X distributed as in (19). Using (20), we can show that the pdf V_t can be reduces to

$$\frac{2\phi\left(\frac{y-\mu_Y}{\sigma_Y}\right)}{\sigma_Y E\left(\mu_{\mathcal{C}_t}(X)\right)}\left((\alpha y+\beta-t)\left(\Phi\left(\frac{t-\alpha y-\beta}{\sigma_{X.Y}}\right)-\Phi\left(\frac{t-1-\alpha y-\beta}{\sigma_{X.Y}}\right)\right)\right.$$
$$\left.-\sigma_{X.Y}\left(\phi\left(\frac{t-\alpha y-\beta}{\sigma_{X.Y}}\right)-\phi\left(\frac{t-1-\alpha y-\beta}{\sigma_{X.Y}}\right)\right)\right) \tag{31}$$

where α, β, and $\sigma_{X.Y}$ are presented in (15) and

$$E\left(\mu_{\mathcal{C}_t}(X)\right) = 2(\mu_X - t)\left(\Phi\left(\frac{t-\mu_X}{\sigma_X}\right)-\Phi\left(\frac{t-1-\mu_X}{\sigma_X}\right)\right)-2\sigma_X\left(\phi\left(\frac{t-\mu_X}{\sigma_X}\right)-\phi\left(\frac{t-1-\mu_X}{\sigma_X}\right)\right). \tag{32}$$

Now, we present a result that obtain a compact form for the mathematical expectation of V_t that can be computed as in (22). This mathematical expectation is in fact a non-linear regression based on fuzzy selection distribution.

Theorem 4.3. *Suppose that the random vector $(X, Y)^T$ has distribution function (15) and consider the membership function (29), then the mathematical expectation of $V_t \stackrel{d}{=} (Y \mid X \in \mathcal{C}_t)$ equals*

$$E(V_t) = \frac{2}{E\left(\mu_{\mathcal{C}_t}(X)\right)}\left(\alpha\Big(\gamma_2(t)-\gamma_2(t-1)\Big)+(\beta-t)\Big(\gamma_3(t)-\gamma_3(t-1)\Big)-\sigma_{X.Y}\Big(\gamma_1(t)-\gamma_1(t-1)\Big)\right),$$

where γ_1, γ_2, and γ_3 are introduced in Theorem 4.2 and $E\left(\mu_{\mathcal{C}_t}(X)\right)$ is presented in (32).

Proof. We can obtain the desirable expectation as $E(V_t) = \int_{x\in\mathbb{R}} y\phi_{FSLCT-N}(y)dy$, where $\phi_{FSLCT-N}(y)$ is computed in (31). Therefore, we have

$$E(V_t) = \int_{-\infty}^{\infty} y\frac{\phi\left(\frac{y-\mu_Y}{\sigma_Y}\right)}{\sigma_Y E\left(\mu_{\mathcal{C}_t}(X)\right)}\left(2(\alpha y+\beta-t)\left(\Phi\left(\frac{t-\alpha y-\beta}{\sigma_{X.Y}}\right)-\Phi\left(\frac{t-1-\alpha y-\beta}{\sigma_{X.Y}}\right)\right)\right.$$
$$\left.-2\sigma_{X.Y}\left(\phi\left(\frac{t-\alpha y-\beta}{\sigma_{X.Y}}\right)-\phi\left(\frac{t-1-\alpha y-\beta}{\sigma_{X.Y}}\right)\right)\right).$$

By Lemma 4.1 we can obtain the results with an argument similar to that one in the proof of Theorem 4.2. □

5 Application

BMI is a measure for human body shape based on an individual's weight and height which is defined as the individual's body mass divided by the square of their height. The formulae universally used in medicine produce a unit of measure of kg/m^2.

For considering the relationship between high blood pressure and BMI, 596 people who met the inclusion criteria of the study were collected in 2012. All subjects (aged 20-64 years) were of Iranian origin and from a central province of the country. They had no organic disease (i.e., liver or kidney disease or diabetes mellitus). Other inclusion criteria were absence of pregnancy or lactation, and no presence of convulsions or its history. Weight, height and blood pressure were measured using standard methods and analyzed for the study stages. Weight was measured while the subjects were wearing light clothing and no shoes. High blood pressure is a chronic medical condition in which the blood pressure in the arteries is elevated. This requires the heart to work harder than normal to circulate blood through the blood vessels. Blood pressure is summarized by two measurements, systolic and diastolic, which depend on whether the heart muscle is contracting (systole) or relaxed between beats (diastole) and equate to a maximum and minimum pressure, respectively.

In this study, we want to consider the relation between high systolic blood pressure and BMI. Normal blood pressure at rest is within the range of 100-140 mmHg systolic (top reading) and 60-90 mmHg diastolic (bottom reading). High blood pressure is said to be present if it is persistently at or above 140/90 mmHg. Here, we describe 140 mmHg as a fuzzy threshold for indicating the high systolic blood pressure and we want to investigate the relationship between high systolic blood pressure and BMI. The scatter plot of BMI versus systolic blood pressure is shown in Figure 1. In this plot, we distinguish the normal blood pressure from high blood pressure, which shown dependency between BMI and blood pressure are different before and after the threshold. We use Energy package in R which provides a multivariate extension to the Shapiro-Wilks test. This normality test proves the sample data has a bivariate normal distribution (P-Value> 0.05) with the following mean vector and variance covariance matrix.

$$\hat{\mu} = \begin{pmatrix} 12.26318 \\ 26.57912 \end{pmatrix}, \quad \hat{\Sigma} = \begin{pmatrix} 4.516108 & 3.242904 \\ 3.242904 & 6.498118 \end{pmatrix}$$

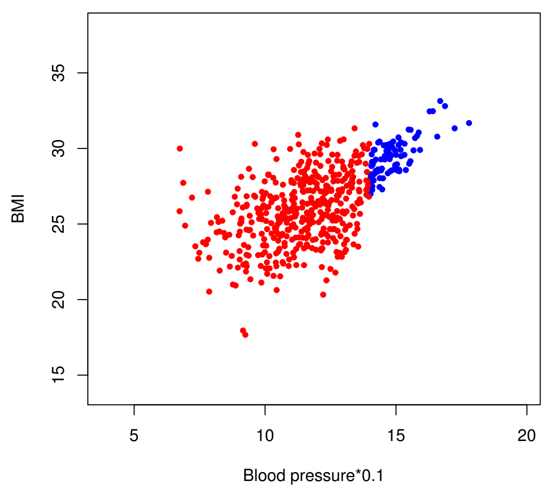

Figure 1: The scatter plot of BMI versus systolic blood pressure, which shows dependency between BMI and blood pressure are differ before and after the threshold.

By (23), we can represent the nonlinear prediction of BMI when the blood pressure is approximately greater than a as

$$\frac{1}{Pr(X \in \mathcal{C}_a)} \Big(2.89775 \big(\gamma_1(a-1) - \gamma_1(a) \big) + 0.4990528 \big(\gamma_2(a-1) - \gamma_2(a) \big) \\ - \big(0.001204 + a \big) \big(\gamma_3(a-1) - \gamma_3(a) \big) + 26.57912 - \gamma_3(a) \Big),$$

where

$$Pr(X \in \mathcal{C}_a) = (13.26318 - a) \big(\Phi(0.4706a - 5.77) - \Phi(0.4706a - 6.24) \big) \\ -2.125114 \big(\phi(0.4706a - 5.77) - \phi(0.4706a - 6.24) \big) + \Phi(5.77 - 0.4706a),$$

$$\gamma_1(a) = (0.575197a + 13.662) \cdot \phi(0.4706a - 5.77),$$

$$\gamma_2(a) = 712.94774 \cdot \Phi(0.4706a - 5.77) - (17.29703 + 0.1287065a) \cdot \phi(0.4706a - 5.77),$$

$$\gamma_3(a) = 26.57912 \cdot \Phi(0.4706a - 5.77) - 1.52599 \cdot \phi(0.4706a - 5.77).$$

Table 1: Comparison results for ordinary with non-linear regression.

Model	R^2	MSE	P-value
Ordinary regression	0.2075	21.487	0.0145
Non-linear regression	0.3621	18.658	0.0015

5.1 Comparison of the proposed non-linear regression with the ordinary regression

In this subsection, we comparison the proposed non-linear regression based on fuzzy selection distribution with the ordinary regression. For this purpose we use the following algorithm to simulate two variables for regression model when the independent variable is a vague variable.
Algorithm

- Generate bivariate random vector (X, Y).
- Replace X with the random variable $Z \sim U(X-1, X+1)$, where U is uniform distribution function.

By the above algorithm, we simulate two variables when the joint distribution of (X, Y) is normal distribution with mean $(1, 2)$ and variance-covariance matrix $\begin{pmatrix} 1 & 0.75 \\ 0.75 & 1 \end{pmatrix}$ and replace X with Z according to the second step of the Algorithm. Then, we fit an ordinary regression and non-linear regression based on fuzzy selection distribution to this data and the results is shown in Table 1. As we can see the non-linear regression model based on fuzzy selection distribution is more efficient than the ordinary regression.

6 Conclusion

Fuzzy selection distributions provides a generalization of the selection distributions that can be obtained by the fuzzy logic. In this paper, we have proposed a fuzzy mechanism to produce a fuzzy selection distribution. Theses results can be used for finding a better relationship between two random vectors. For example in epidemiology it is a useful method to consider relation between behavioral and physiological variables, when we want to investigate them in fuzzy environment. We can also fuzzy selection distribution to extend the classical measures of relationship between two vectors where one of them belong to a certain rang, especially in quality control, mathematical insurance, and economics.

References

[1] Ajani, U., Dunbar, S., Ford, E., Mokdad, A., and Mensah, G. (2005). Sodium intake among people with normal and high blood pressure. *Am. J. Prev. Med.* 29 (5 suppl. 1), 63–67.
[2] Arellano-Valle, R. B., Branco, M. D., and Genton, M. G. (2006). A unified view on skewed distributions arising from selections. *Canad. J. Stat.* 34, 581–601.
[3] Arellano-Valle, R. B. and Azzalini, A. (2006). On the unification of families of skew-normal distributions. *Scand. J. Stat.* 33, 561–574.
[4] Arellano-Valle, R. B. and Genton, M. G. (2005). On fundamental skew distribution. *J. Multivar. Anal.* 96, 93–116.
[5] Arellano-Valle, R. B. and Genton, M. G. (2010). Multivariate unified skew-elliptical distributions. *Chilean J. Stat.* 1, 17–33.
[6] Azzalini, A. (1985). A class of distributions which includes the normal ones. *Scand. J. Stat.* 12, 171–178.
[7] Azzalini, A. (2005). The skew-normal distribution and related multivariate families. *Scand. J. Stat.* 32, 159–188.
[8] Azzalini, A. and Dalla Valle, A. (1996). The multivariate skew-normal distribution, *Biometrika* 83, 715–726.
[9] Billingsley, P. (1995). *Probability and Measure*. John Wiley & Sons, New York.
[10] Dubois, D. and Prade, H. (1980) *Fuzzy Sets and Systems: Theory and Applications*, Academic Press, New York.

[11] Flores-Huerta, S., Klünder-Klünder, M., Reyes de la Cruz, L. and Santos J.I. (2009). Increase in body mass index and waist circumference is associated with high blood pressure in children and adolescents in Mexico City. *Arch. Med. Res.* 40(3), 208–215.
[12] Genton, M.G. (2004). *Skew-Elliptical Distributions and Their Applications: A Journey Beyond Normality*. Chapman & Hall/CRC, Boca Raton, FL.
[13] Kaufman, J, Asuzu, M, Mufunda, J., Forrester, T., Wilks, R., Luke, A., Long, A, and Cooper, R. (1997). Relationship between blood pressure and body mass index in lean populations. *Hypertension* 30, 1511–1516.
[14] Lupton, S., Chiu, C., Lujic, S., Hennessy, A., and Lind J. (2013). Association between parity and breastfeeding with maternal high blood pressure. *Am. J. Obstet. Gynecol.* 208(6), 454.e1–e7.
[15] Pourmousa, R., Mashinchi, M. (2012). An interpretation of skew-elliptical distributions in terms of fuzzy events. *Scientia Iranica E* 19(6), 1870–1875.
[16] Jerez, S., and Coviello, A. (1998). Alcohol Drinking and blood pressure among adolescents. *Alcohol* 16(1), 1–5.
[17] Ueshima, H., Ozawa, H., Baba, H., Nakamoto, Y., Omae, T., Shimamotos, T., and Komachi, Y. (1992). Alcohol drinking and high blood pressure: data from a 1980 national cardiovascular survey of Japan. *J. Clin. Epidemiol.* 45(6), 667–673.
[18] Zadeh, L.A. (1965). Fuzzy sets. *Information Control* 8, 338–353.
[19] Zadeh, L.A. (1968). Probability measures of fuzzy events, *J. Math. Anal. Appl.* 23, 421–427.
[20] Ziaei, S., Janghorban, R., Shariatdoust, S, and Faghihzadeh S. (2008). The effects of iron supplementation on serum copper and zinc levels in pregnant women with high-normal hemoglobin, *Int. J. Gynaecol. Obstet.* 100(2), 133–135.