PROCEEDINGS OF THE YEREVAN STATE UNIVERSITY

Physical and Mathematical Sciences 2017, **51**(1), p. 8–12

Mathematics

FUZZY BAYESIAN INFERENCES

V. G. BARDAKHCHYAN *

Chair of Actuarial Mathematics and Risk Management YSU, Armenia

We consider generalization of Bayesian theorem for the cases with fuzzy observations. It will be considered one of the forms of fuzzy Bayesian theorem introduced by Viertl. One can observe that it cannot be used as a generalization of ordinary Bayesian theorem with the form of likelihood function. We introduce other types of likelihood function satisfying some definite conditions.

MSC2010: Primary 28E10; Secondary 60A86, 62F15.

***Keywords*:** fuzzy statistics, Bayesian theorem, fuzzy random variables.

Introduction. The fuzzy measure theory has generated a lot of new concepts, such as fuzzy random variables, their distributions and density functions. It also arised new problems, like statistics concerned with fuzzy observations, hypothesis testing for such data, modeling with such data (for example, fitting with linear fuzzy regressions) and so on (see, f.e., [1]). Our study is concerned with fuzzy Bayesian analysis. First this question was considered by Viertl in one of his reports. He emphasized the necessity of fuzzy Bayesian theorem and he described the structure of the theory and the ways to make fuzzy analog of Bayesian theorem. His approach can be found in [2, 3]. In these studies the theorem and related concepts and properties are described. .

In his definition the fuzzy Bayesian theorem has the following form:

$$\underline{\pi}_\delta(\theta|x_1,x_2,...,x_n) = \frac{2\underline{\pi}_\delta(\theta)l(\theta,x_1,x_2,...,x_n)}{\int(\underline{\pi}_\delta(\theta)+\overline{\pi}_\delta(\theta))l(\theta,x_1,x_2,...,x_n)d\theta}, \tag{1}$$

$$\overline{\pi}_\delta(\theta|x_1,x_2,...,x_n) = \frac{2\overline{\pi}_\delta(\theta)l(\theta,x_1,x_2,...,x_n)}{\int(\underline{\pi}_\delta(\theta)+\overline{\pi}_\delta(\theta))l(\theta,x_1,x_2,...,x_n)d\theta}. \tag{2}$$

Where $\underline{\pi}_\delta(\theta|x_1,x_2,...,x_n)$ and $\overline{\pi}_\delta(\theta|x_1,x_2,...,x_n)$ stand for lower and upper bounds of the a posterior fuzzy density of δ-level, $\underline{\pi}_\delta(\theta)$ and $\overline{\pi}_\delta(\theta)$ are those for respective prior distribution, $\delta\in[0,1]$ is the level of confidence, i.e. the level of belonging in the theory of fuzzy sets, $l(\theta,x_1,x_2,...,x_n)$ is the likelihood function (LF). Most of the questions arising in this field is concerned with the definition of this LF, when the observations $x_1,...,x_n$ are fuzzy.

For these variants of definition of posterior fuzzy densities it was proved that sequential usage of the theorem and the using of all collected data at once gives the same result for the posterior distributions. The proof can be found in [2–4]. This property is essential, as if we could find the distribution, which is self conjugate in the sense of posterior and prior

* E-mail: vardanbardakchyan@gmail.com

distributions for one observation (meaning that the posterior and prior distributions are of the same families), it would be self-conjugate for any size of samples.

Viertl's LF has the following form

$$l(\theta,x_1,x_2,...,x_n)=\prod_i\left(\int_{l_i}^{r_i}\xi_i(x)f(x|\theta)dx\right), \tag{3}$$

here $\xi_i(x)$ is the membership function of the fuzzy observation x_i, $[l_i,r_i]$ is its support, and $f(x|\theta)$ is the density of the conditional distribution see for more details [2,4]. Viertl has given a further generalization of the theorem, taking LF to be fuzzy [2].

The problem we consider here is dealing with the normal distribution in the fuzzy case. The normal distribution has the above mentioned self-conjugate property. For the fuzzy version this is not the case in reality, because a normally distributed variable will note remain normal for fuzzy triangle observations. Moreover, if we consider more than one observation the precise calculation of new parameters becomes extremely hard.

In the paper two other versions of fuzzy Bayesian theorems are introduced, more precisely for LF, since the above stated version has some flaws. Some related questions are considered in [5,6].

Bayesian Theorem with Fuzzy Triangular Observations. Here only the fuzzy triangular number for observations are considered and more explicit formulas for the theorem are given. Let's start with the formula making flaws of the above mentioned version of LF explicit.

Proposition 1. If our observation in Eq. (3) is a triangular fuzzy number, i.e. $\xi_i(x)=\begin{cases}(x-l_i)/(m_i-l_i), & \text{if } x\in[l_i,m_i]\\(r_i-x)/(r_i-m_i), & \text{if } x\in[m_i,r_i]\end{cases}$, then the Eq. (3) takes the following form:

$$l(\theta,x_1,x_2,...,x_n)=\prod_i\left(\frac{1}{r_i-m_i}\int_{m_i}^{r_i}F(x|\theta)dx-\frac{1}{m_i-l_i}\int_{l_i}^{m_i}F(x|\theta)dx\right). \tag{4}$$

Proof. Consideration of one observation will be enough, and after simple operations (integrating by parts) we get

$$\begin{aligned}\int_{l_i}^{r_i}\xi_i(x)f(x|\theta)dx&=\int_{l_i}^{m_i}\frac{x-l_i}{m_i-l_i}f(x|\theta)dx+\int_{m_i}^{r_i}\frac{r_i-x}{r_i-m_i}f(x|\theta)dx\\&=F(m_i|\theta)-\frac{1}{m_i-l_i}\int_{l_i}^{m_i}F(x|\theta)dx-F(m_i|\theta)+\frac{1}{r_i-m_i}\int_{m_i}^{r_i}F(x|\theta)dx\\&=\frac{1}{r_i-m_i}\int_{m_i}^{r_i}F(x|\theta)dx-\frac{1}{m_i-l_i}\int_{l_i}^{m_i}F(x|\theta)dx.\end{aligned}$$ □

The last expression means that the likelihood function with fuzzy triangular observation comes to the difference of the average values of cumulative distribution function on the $[m_i,r_i]$ and on the $[l_i,m_i]$.

Remark. One can observe, that if $l_i\to r_i$ the function vanishes for the continuous distributions. Therefore this cannot be taken as a generalization of ordinary LF for continuous random variable, other extensions as LF should be taken for fuzzy observations.

Generally the following proposition is true.

Proposition 2. Let a prior distribution be collection of fuzzy triangular numbers for each value of arguments and for $\delta=1$ have a normal distribution. If the observations are crisp (non-fuzzy) from normal distribution, then the posterior fuzzy distribution is normal at top, i.e. $\delta=1$ when we use ordinary LF for crisp data.

For this simple argument, one should just verify that $\underline{\pi}_1(\theta) = \overline{\pi}_1(\theta) = f(\theta)$.

However, as one can see, this fails to be in the case when the observations are fuzzy. The situation in that case can be stated by the following theorem.

Theorem. For the prior normal distribution ($N(\mu_0, \sigma_0^2)$) and the fuzzy triangular observations from the normal distribution ($N(\theta, \sigma^2)$), the posterior distribution is not normal. For one observation the new estimate of the mean (location) parameter is

$$\hat{\theta} = \mu_0 \sigma^2/(\sigma^2 + \sigma_0^2) + \sigma_0^2 B/((\sigma^2 + \sigma_0^2)A),$$

where A and B are given below (see Eqs. (6), (8)).

Proof. So we have that for every $\delta \in [0,\ 1]$

$$g(\theta) = \underline{\pi}_\delta(\theta) = \overline{\pi}_\delta(\theta) = \frac{1}{\sigma_0\sqrt{2\pi}} e^{-\frac{(\theta-\mu_0)^2}{2\sigma_0^2}}, \quad f(x|\theta) = \frac{1}{\sigma\sqrt{2\pi}} e^{-\frac{(x-\theta)^2}{2\sigma^2}},$$

$$\xi(x) = \begin{cases} (x-l)/(m-l), & l \le x \le m, \\ (r-x)/(r-m), & m \le x \le r. \end{cases}$$

The new estimate will be

$$\hat{\theta} = \int_\Theta \theta g(\theta) l(x,\theta) d\theta \Big/ \int_\Theta g(\theta) l(x,\theta) d\theta\,. \tag{5}$$

Let's start with the denominator,

$$\begin{aligned} \int_\Theta g(\theta) l(x,\theta) d\theta &= \int_{-\infty}^{\infty} \frac{1}{\sigma_0\sqrt{2\pi}} e^{-\frac{(\theta-\mu_0)^2}{2\sigma_0^2}} \left(\int_{-\infty}^{\infty} \xi(x) \frac{1}{\sigma\sqrt{2\pi}} e^{-\frac{(x-\theta)^2}{2\sigma^2}} dx \right) d\theta \\ &= \frac{1}{2\sigma_0\sigma\pi} \int_{-\infty}^{\infty} \xi(x) \left(\int_{-\infty}^{\infty} e^{-\frac{(\theta-\mu_0)^2}{2\sigma_0^2} - \frac{(x-\theta)^2}{2\sigma^2}} d\theta \right) dx. \end{aligned}$$

Taking $\alpha = \dfrac{\sigma^2}{\sigma^2 + \sigma_0^2}$ and $\beta^2 = \dfrac{1}{\sigma^2 + \sigma_0^2}$, we will have

$$\begin{aligned} \int_\Theta g(\theta) l(x,\theta) d\theta &= \frac{1}{2\sigma_0\sigma\pi} \int_{-\infty}^{\infty} \xi(x) e^{-\frac{((x-\mu_0)\beta)^2}{2}} \sqrt{2\pi\alpha(1-\alpha)(\sigma^2+\sigma_0^2)} dx \\ &= \frac{\beta}{\sqrt{2\pi}} \int_{-\infty}^{\infty} \xi(x) e^{-\frac{((x-\mu_0)\beta)^2}{2}} dx. \end{aligned}$$

$$\begin{aligned} \int_{-\infty}^{\infty} g(\theta) l(x,\theta) d\theta = & \frac{1}{(m-l)\beta} \Big(\phi\big((l-\mu_0)\beta\big) - \phi\big((m-\mu_0)\beta\big) \Big) \\ & - \frac{1}{(r-m)\beta} \Big(\phi\big((m-\mu_0)\beta\big) - \phi\big((r-\mu_0)\beta\big) \Big) \\ & + \frac{\mu_0 - l}{m-l} \Big(\Phi\big((m-\mu_0)\beta\big) - \Phi\big((l-\mu_0)\beta\big) \Big) \\ & + \frac{r-\mu_0}{r-m} \Big(\Phi\big((r-\mu_0)\beta\big) - \Phi\big((m-\mu_0)\beta\big) \Big) =: A, \end{aligned} \tag{6}$$

where ϕ and Φ are density and cumulative distribution functions of standard normal distribution respectively.

Taking into account Eq. (6), we will have for the numerator:

$$\int_\Theta \theta g(\theta) l(x,\theta) d\theta = \mu_0 \alpha A + (1-\alpha) B, \tag{7}$$

where B is the following:

$$
\begin{aligned}
B :=& \frac{r}{(r-m)\beta}\Big(\phi\big((m-\mu_0)\beta\big)-\phi\big((r-\mu_0)\beta\big)\Big)\\
&+\ \mu_0\frac{r}{r-m}\Big(\Phi\big((r-\mu_0)\beta\big)-\Phi\big((m-\mu_0)\beta\big)\Big)\\
&-\ \frac{l}{(m-l)\beta}\Big(\phi\big((l-\mu_0)\beta\big)\ -\ \phi\big((m-\mu_0)\beta\big)\Big)\\
&-\ \mu_0\frac{l}{m-l}\ \Big(\Phi\big((m-\mu_0\beta\big)\ -\ \Phi\big((l-\mu_0)\beta\big)\Big)\\
&+\ \frac{2\mu_0}{(m-l)\beta}\Big(\phi\big((l-\mu_0)\beta\big)\ -\ \phi\big((m-\mu_0)\beta\big)\Big)\\
&+\ \frac{\mu_0^2}{m-l}\ \Big(\Phi\big((m-\mu_0)\beta\big)\ -\ \Phi\big((l-\mu_0)\beta\big)\Big)\\
&+\frac{1}{(m-l)\beta^2}\Big(\Phi\big((m-\mu_0)\beta\big)-\Phi\big((l-\mu_0)\beta\big)+\\
&+\phi\big((l-\mu_0)\beta\big)(l-\mu_0)\beta-\phi\big((m-\mu_0)\beta\big)(m-\mu_0)\beta\Big)\\
&-\frac{2\mu_0}{(r-m)\beta}\Big(\phi\big((m-\mu_0)\beta\big)-\phi\big((r-\mu_0)\beta\big)\Big)\\
&-\frac{\mu_0^2}{r-m}\Big(\Phi\big((r-\mu_0)\beta\big)-\Phi\big((m-\mu_0)\beta\big)\Big)\\
&-\frac{1}{(r-m)\beta^2}\Big(\Phi\big((r-\mu_0)\beta\big)-\Phi\big((m-\mu_0)\beta\big)+\\
&+\phi\big((m-\mu_0)\beta\big)(m-\mu_0)\beta-\phi\big((r-\mu_0)\beta\big)(r-\mu_0)\beta\Big).
\end{aligned}
\tag{8}
$$

In the expression the following equations are used (see [7], p. 393):
$\int \phi(x)dx = \Phi(x)+C;\ \int x\phi(x)dx = -\phi(x)+C;\ \int x^2\phi(x)dx = \Phi(x)-x\phi(x)+C.$
So finally we get

$$\hat{\theta} = \alpha\mu_0 + (1-\alpha)\frac{B}{A}. \tag{9}$$

Obviously further computations cannot be done directly, since the new θ is not of normal distribution.

Other Forms for LF. As we stated in Remark 2, the used LF cannot be taken as a generalization of ordinary LF. So let's look for other possible approaches. One of them can be stated by formula

$$l(\theta, x_1, x_2, ..., x_n) = \prod_i f(DF(x_i)|\theta), \tag{10}$$

where $DF(\cdot)$ denotes some defuzzification for any fuzzy input, generally not depending on exogenous parameter θ and if so, then we will have the property for normal function as in crisp case. That is for any fuzzy observations of normal distribution prior normal distribution will remain normal. This is because the problem with the LF used by Viertl is that this generalization does not express probability.

By another approach we evaluate the intervals of probabilities for each level for a given distribution in LF. That is as new fuzzy numbers we will have α-cuts the following intervals $y_\alpha = [\min_{x\in x_\alpha} f(x|\theta),\ \max_{x\in x_\alpha} f(x|\theta)]$, where x_α stands for α-cut of the original fuzzy observation. So we will have maximal and minimal probabilities for each level of membership. And as new fuzzy number we take $\eta(y) = \max\{\alpha | y \in y_\alpha\}$. Note that we evaluate the maximal and minimal probabilities depending as well on θ. As a LF for this case we take

$$l(\theta, x_1, x_2, ..., x_n) = \prod_i DF(y_i). \tag{11}$$

This technique is very alike (but differs essentialy) to the one used for the construction of Sugeno integral (see [8]). Note that the y_i-es express probabilities, so this is much more likely to be more convenient generalization of LF.

However, in this case depending on the defuzzification technique that we use, the property for normal distribution can fail to be fulfilled. For example, taking as defuzzification $DF(y) = \int \eta(y) y dy$, one can see that the normal does not remain normal.

For simplicity let's examine a symmetrical triangle $\left(\text{i.e. } m = \dfrac{l+r}{2}\right)$. We will have the following cases:

Case 1. $\theta \in (-\infty, l]$; *Case 2.* $\theta \in [r, \infty)$; *Case 3.* $\theta \in (l, m]$; *Case 4.* $\theta \in [m, r)$.

For example, in the Case 1 the following form for η is obtained:

$$\eta(y) = \begin{cases} \dfrac{r - \sigma(-2\ln(y\sigma\sqrt{2\pi}))^{1/2} - \theta}{r-m}, & y \in \left(\dfrac{1}{\sigma\sqrt{2\pi}} e^{-\frac{(r-\theta)^2}{2\sigma^2}}, \dfrac{1}{\sigma\sqrt{2\pi}} e^{-\frac{(m-\theta)^2}{2\sigma^2}}\right], \\ \dfrac{\sigma(-2\ln(y\sigma\sqrt{2\pi}))^{1/2} + \theta - l}{m-l}, & y \in \left[\dfrac{1}{\sigma\sqrt{2\pi}} e^{-\frac{(m-\theta)^2}{2\sigma^2}}, \dfrac{1}{\sigma\sqrt{2\pi}} e^{-\frac{(l-\theta)^2}{2\sigma^2}}\right), \end{cases}$$

and for this case, so for the all others, the distribution does not remain normal.

$$DF(y) = \int \eta(y) y dy = \int_{m-\theta}^{r-\theta} \frac{r-t-\theta}{r-m} \cdot \frac{1}{2\pi\sigma^2} \cdot \frac{t}{\sigma^2} e^{-\frac{t^2}{2\sigma^2}} dt + \\ + \int_{l-\theta}^{m-\theta} \frac{t+\theta-l}{m-l} \cdot \frac{1}{2\pi\sigma^2} \cdot \frac{t}{\sigma^2} e^{-\frac{t^2}{2\sigma^2}} dt.$$

Since θ is under the integral, $\Phi + \phi$ form can be obtained, from which obviously we will not have a normal distribution. However, for some other types of defuzzification we will have the desired property (for example, if we take the tops).

The author would like to express his gratitude to the reviewer, for a fruitful discussion and important comments made.

Received 19.10.2016

R E F E R E N C E S

1. **Klirl J.G., Yuan B.** Fuzzy Sets and Fuzzy Logic: Theory and Applications. Prentice Hall PTR, 1995.
2. **Viertl R.** Statistical Methods for Fuzzy Data. Chichester: Wiley, 2011.
3. **Viertl R., Sunanta O.** Fuzzy Bayesian Inference. // METRON, 2013, v. 71, № 3, p. 207–216.
4. **Buckley J.** Fuzzy Probability and Statistics. Berlin, Heidleberg: Springer–Verlag, 2006.
5. **Yang C.C.** Fuzzy Bayesian Inference. // Proceeding of Systems, Man and Cybernetics. IEEE International Conference, 1997, v. 3, p. 2707–2712.
6. **Osoba O., Miataim S., Kosko B.** Bayesian Inference with Adaptive Fuzzy Priors and Likelihoods. // IEEE Transactions on Systems, Man and Cybernetics, 2011, v. 41, № 5, p. 1183–1197.
7. **Owen D.B.** A Table of Normal Integrals. // Communication in Statistics – Simulation and Computation, 1980, v. 9, № 4, p. 389–419.
8. **Sugeno M.** Theory of Fuzzy Integrals and Its Applications. PhD Thesis. Tokyo Institute of Technology, 1974.