Fuzzy Bayesian Inference

Christopher C. Yang
Department of Computer Science
The University of Hong Kong
Email: yang@cs.hku.hk

Abstract

Bayesian methods provide formalism for reasoning about partial beliefs under conditions of uncertainty. Given a set of exhaustive and mutually exclusive hypotheses, one can compute the probability of a hypothesis for given evidence using the Bayesian inversion formula. In the Bayesian inference, the evidence could be a single atomic proposition or multi-valued. For multi-valued evidence, these values could be discrete, continuous, or fuzzy. For continuous-valued evidence, the density functions used in the Bayesian inference are difficult to be determined in many practical situations. Complicated laboratory testing and advance statistical techniques are required to estimate the parameters of the assumed type of distribution. Using the proposed fuzzy Bayesian approach, formulation is derived to estimate the density function from the conditional probabilities of the fuzzy-supported values. It avoids the complicated testing and analysis, and it does not require the assumption of a particular type of distribution. The estimated density function in our approach is proved to conform to two axioms in the theorem of the probability. Example is provided in the paper.

1. Introduction

Bayesian theorem is an effective tool for reasoning under the condition of uncertainty. Evidences are multi-valued in some situations and the values may be discrete, continuous, and fuzzy. Propositions are given numerical parameters representing their degree of beliefs under some body of knowledge, these parameters are then combined and manipulated based on the rules of probability theory. P(H|e) represents the subjective belief in the hypothesis, H, given the knowledge of the evidence, e.

To compute the probability of a single evidence, e, by conditioning e on a set of exhaustive and mutually exclusive hypotheses H_j (j = 1, 2, ..., m), we use the following formula:

$$P(e) = \sum_{j=1}^{m} P(e \mid H_j) P(H_j) \tag{1}$$

This formalism states the belief in the evidence, e, is a weighted sum over the beliefs in all the distinct ways that e might be realized given the knowledge of the hypotheses, H_j.

The Bayesian inversion formula is utilized to compute the *posterior probability* of a hypothesis H_j upon the evidence e, $P(H_j|e)$, by multiplying the *prior probability*, $P(H_j)$, to the *likelihood*, $P(e|H_j)$, and divided by P(e):

$$P(H_j \mid e) = \frac{P(e \mid H_j) P(H_j)}{P(e)} \tag{2}$$

P(e) is computed using Equation (1). The Bayesian inversion formula is regarded as a normative rule for updating beliefs in response to evidence.

In some situations, the evidence is multi-valued instead of a single atomic proposition. The values could be discrete, continuous, or fuzzy. For example, length may have multiple discrete values (length (cm) = {100, 110, 120, ..., 200}), or continuous values (length = [100,200]), or fuzzy values (length = {short, medium, long}).

1.1 Discrete Values or Fuzzy Values

Given that e has n discrete values or fuzzy values, $e = e_1, e_2, \ldots, e_n$, and there are m hypotheses,

H_1, H_2, … , H_m, the posterior probability of H_j with evidence e_i is computed as follow:

$$P(H_j|e_i) = \frac{P(e_i|H_j)P(H_j)}{[P(e_1|H_1) \;\; P(e_1|H_2) \dots \;\; P(e_1|H_m)][P(H_1) \;\; P(H_2) \dots \;\; P(H_m)]^T} \tag{3}$$

1.2 Continuous Values

Given that e has continuous values, e =[a,b], and there are m hypotheses, H_1, H_2, … , H_m, the posterior probability of H_j can be computed as follow:

$$P(H_j \mid e) = \frac{f(e \mid H_j)P(H_j)}{\sum_{j=1}^{m} f(e \mid H_j)P(H_j)} \tag{4}$$

where $f(e|H_j)$ is the likelihood density and e is a value between a and b .

In this paper, a methodology is presented to estimate the likelihood density function of the continuous valued evidence from the likelihood probabilities of the fuzzy valued evidences. It is shown that the estimated density functions conform to two axioms of the probability theory. Using the estimated likelihood density functions and the prior probabilities, we can determine the posterior probabilities for a given continuous valued evidence.

2. Related Work in Fuzzy Bayesian Approach

Fuzzy Bayesian approach has been adopted to enhance the probability updating process with fuzzy evidences by utilizing the conditional probability densities and the membership functions of the evidence's values. This approach has been widely applied in structural reliability to access the safety of the constructed projects [1,3].

A fuzzy set is characterized by a membership function, $\mu_e(x)$, which assigns each value of x for the evidence, e, a grade of membership ranging from 0 to 1. For example, an evidence, e, has n fuzzy values, $e_1, e_2, \dots, e_n$.

Given that the value of the evidence is e_i, the likelihood $P(e_i|H_j)$ is computed in [1,3]:

$$P(e_i \mid H_j) = \int_{x \in e_i} \mu_{e_i}(x) f(x \mid H_j) dx \tag{5}$$

where $f(x|H_j)$ is the likelihood density function evaluated at value x given the hypothesis, H_j. We can compute the posterior probability using Equations (1), (2), and (5) as follows:

$$P(H_j \mid e_i) = \frac{\int_{x \in e_i} \mu_{e_i}(x) f(x \mid H_j) dx P(H_j)}{\sum_{k=1}^{m} \left(\int_{x \in e_i} \mu_{e_i}(x) f(x \mid H_k) dx P(H_k) \right)}$$

One can use this approach to determine the likelihood probability for a given fuzzy value from the likelihood density function and then determine the posterior probability. However, the likelihood density function is not easy to be determined. In many of its applications [1,3], the density functions are approximated as a particular type of distribution such as Guassian and Weibull, and the parameters of the approximated distributions are estimated by laboratory testing and statistical methodology. These estimation is complicated and time consuming. Moreover, it is a waste of effort to determine the likelihood probability for the fuzzy valued evidence once we have the likelihood density function for the continuous valued evidence. It is more appropriate to determine the likelihood density function for the continuous valued evidence from the likelihood probability for the fuzzy valued evidence. In the next section, the approach to *estimate* the likelihood density function of the continuous valued evidence from their likelihood probability of the fuzzy valued evidence is introduced and these likelihood probabilities are then utilized to determine the posterior probabilities.

3. Fuzzy Bayesian Inference from Fuzzy Valued Evidence to Continuous Valued Evidence

Given the likelihood probability of the fuzzy valued evidence, an approach to estimate the likelihood density function of the corresponding continuous evidence is developed. The likelihood for fuzzy valued evidence is observed from professional experiences. Each fuzzy value, e_i, covers a range of the continuous values ,and the size of the range is $W(e_i)$. The grade of membership for a particular value in the fuzzy value is measured by the membership functions. Using the membership functions and the likelihood probabilities of the fuzzy values, we can estimate the likelihood density functions as follow:

$$f^*(e \mid H_j) = c\sum_{i=1}^{n} \frac{\mu_{e_i}(e)}{W(e_i)} P(e_i \mid H_j) \tag{6}$$

where

$$c = \frac{W(e_i)}{\int_{e \in e_i} \mu_{e_i}(e)de}$$

In this formulation, it is imposed that the size of the interval divided by the area of the membership function for each fuzzy value is the same and is equal to c.

The estimated likelihood for continuous valued evidence is computed as the weighted sum of the likelihood of the fuzzy valued evidence. The weight is proportional to the corresponding degree of belonging obtained by the membership function and is inversely proportional to the corresponding size of the interval of fuzzy value.

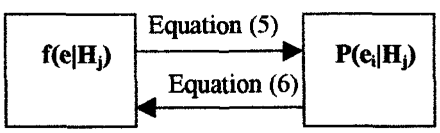

Figure 1. Computing the likelihood probability for fuzzy valued evidence by Equation (5) and computing the likelihood density function for continuous valued evidence by the proposed approach using Equation (6).

It is also proved that the likelihood density function developed in Equation (6) conform to the axioms of probability theory. For any density function f(x), the first axiom requires

$$f(x) \geq 0,$$

and the second axiom requires

$$\int_{-\infty}^{+\infty} f(x)dx = 1.$$

Proof of Axiom 1

$$P(e_i \mid H_j) \geq 0,$$
$$\mu_{e_i}(e) \geq 0,$$
$$W(e_i) \geq 0,$$

Therefore,

$$c \geq 0,$$

and

$$f^*(e \mid H_j) \geq 0$$

Proof of Axiom 2

$$\int_{-\infty}^{+\infty} f^*(e \mid H_j)de$$
$$= \int_{-\infty}^{+\infty} c\sum_{i=1}^{n}\left(\frac{\mu_{e_i}(e)}{W(e_i)} P(e_i \mid H_j)\right)de$$
$$= c\sum_{i=1}^{n}\left(\int_{-\infty}^{+\infty} \frac{\mu_{e_i}(e)}{W(e_i)} P(e_i \mid H_j)de\right)$$
$$= c\sum_{i=1}^{n}\left(\int_{-\infty}^{+\infty} \frac{\mu_{e_i}(e)}{W(e_i)} de\, P(e_i \mid H_j)\right)$$
$$= c\sum_{i=1}^{n}\left(\frac{1}{c} P(e_i \mid H_j)\right)$$
$$= \sum_{i=1}^{n} P(e_i \mid H_j)$$
$$= 1$$

Using the estimated likelihood for the continuous valued evidence, one can compute the posterior probability using Equation (4) and (6).

$$P(H_j \mid e) = \frac{f^*(e \mid H_j)P(H_j)}{\sum_{j=1}^{m} f^*(e \mid H_j)P(H_j)}$$

4. Example

In this section, a steak-cooking example is utilized to illustrate the fuzzy Bayesian approach presented in Section 3. Let's assume that there are three hypotheses, (i) the steak is burnt (B), (ii) the steak is well-done (W), and (iii) the steak is not done yet (N). There are five fuzzy values for the evidence, temperature of the stove, low (L), medium low (ML), medium (M), medium high (MH), and high (H). Using the formulation in Section 3, one can determine the posterior probability, e.g. the probability of the steak is burnt given the temperature of the stove is high.

The membership functions for the fuzzy values are as follow:

$$\mu_L(e) = -\frac{1}{50}e + 5 \qquad 200 \le e \le 250$$

$$\mu_{ML}(e) = \begin{cases} \frac{1}{50}e - 4 & 200 \le e \le 250 \\ -\frac{1}{50}e + 6 & 250 \le e \le 300 \end{cases}$$

$$\mu_M(e) = \begin{cases} \frac{1}{50}e - 5 & 250 \le e \le 300 \\ -\frac{1}{50}e + 7 & 300 \le e \le 350 \end{cases}$$

$$\mu_{MH}(e) = \begin{cases} \frac{1}{50}e - 6 & 300 \le e \le 350 \\ -\frac{1}{50}e + 8 & 350 \le e \le 400 \end{cases}$$

$$\mu_H(e) = \frac{1}{50}e - 7 \qquad 350 \le e \le 400$$

Table 1 provides the likelihood probabilities, $P(e_i|H_j)$ such as P(L|B), for the fuzzy values and Table 2 provides the prior probabilities of the hypotheses, $P(H_j)$ such as P(B).

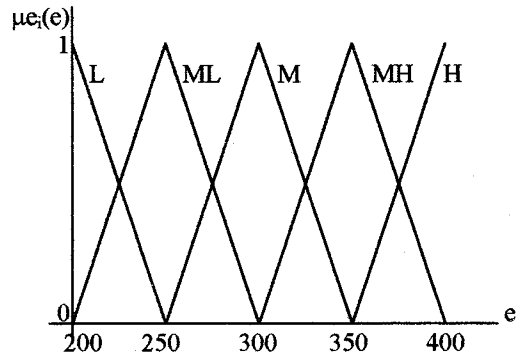

Figure 2. Membership functions for the values of the evidence are low, medium low, medium, medium high, and high.

Table 1. The likelihood probabilities of the steak-cooking example.

	L	ML	M	MH	H
B	0.00	0.05	0.10	0.20	0.65
W	0.05	0.05	0.40	0.40	0.20
N	0.60	0.20	0.10	0.05	0.05

Table 2. The prior probabilities of the steak-cooking example.

P(B)	0.20
P(W)	0.60
P(N)	0.20

Using the membership functions and the likelihood probabilities in Table 1, we can compute the likelihood density functions using Equation (6). Let's assume e = 283.

$$f^*(283 \mid B) = c\left(\frac{0.00}{50}\mu_L(283) + \frac{0.05}{100}\mu_{ML}(283) + \frac{0.10}{100}\mu_M(283) + \frac{0.20}{100}\mu_{MH}(283) + \frac{0.65}{50}\mu_H(283)\right)$$

Since c = 2, $\mu_L(283) = 0$, $\mu_{ML}(283) = 0.34$, $\mu_M(283) = 0.66$, $\mu_{MH}(283) = 0$, and $\mu_H(283) = 0$,

$f^*(283|B) = 0.00166$.

Similarly,

$$f^*(238\,|\,W) = 2\left(\frac{0.05}{100}0.34 + \frac{0.4}{100}0.66\right)$$
$$= 0.00528$$

$$f^*(238\,|\,N) = 2\left(\frac{0.20}{100}0.34 + \frac{0.10}{100}0.66\right)$$
$$= 0.00268$$

We find that the probability of the temperature of the stove is 283 given that the steak is well done is the highest (=0.00528), and the probability of the temperature is 283 given that the steak is burnt is the lowest (=0.00166). The probability of the temperature is 283 given that the steak is not done is the second highest (=0.00268). The probabilities are small (less than 0.01) because the evidence (temperature) has continuous value.

Using the estimated likelihood density functions and the prior probabilities, we can determine the posterior probability.

$$P(B\,|\,283) = \frac{f^*(283\,|\,B)P(B)}{f^*(283\,|\,B)P(B) + f^*(283\,|\,W)P(W) + f^*(283\,|\,N)P(N)}$$
$$= \frac{(0.00166)(0.20)}{(0.00166)(0.20) + (0.00528)(0.60) + (0.00268)(0.20)}$$
$$= \frac{0.000332}{0.004036}$$
$$= 0.082$$

Similarly,

$$P(W\,|\,283) = \frac{(0.00528)(0.60)}{(0.00166)(0.20) + (0.00528)(0.60) + (0.00268)(0.20)}$$
$$= 0.785$$

$$P(N\,|\,283) = \frac{(0.00268)(0.20)}{(0.00166)(0.20) + (0.00528)(0.60) + (0.00268)(0.20)}$$
$$= 0.133$$

As a result, given the temperature is 283, the probability of well done for the steak is the highest which is 0.785 and the probability of burnt is the lowest which is 0.082, the probability of not done is 0.133. (The total of these probabilities is 1.00.) We find that the probability of the steak is well done is almost ten times of the probability of the steak is burnt given that the temperature of the stove is 283.

5. Conclusion

The fuzzy Bayesian inference presented in this paper provides a mechanism for determining the posterior probabilities given a value of the continuous evidence, provided that the fuzzy likelihood probabilities and the prior probabilities are known. In the previous work, mechanism has been developed to determine posterior probabilities given the fuzzy value of the evidence. However, the likelihood density functions are approximated by tedious and complicated laboratory work. Example is presented to illustrate the mechanism. Moreover, proofs are provided to show that the formulation conform to the axioms of theory of probability.

References

[1] Chou, K. C., and Yuan, J., "Fuzzy-Bayesian Approach to Reliability of Existing Structures," *Journal of Structural Engineering*, vol.119, no.11, November, 1993.

[2] Furthsirth-Schnater, S., "On Fuzzy Bayesian Inferences," *Fuzzy Sets and Systems*, vol.60, 1993.

[3] Itoh, S., and Itagaki, H., "Applications of Fuzzy-Bayesian Analysis to Structural Reliability," *Proceedings of ICOSSAR '89, the 5th International Conference on Structural Safety and Reliability*, San Francisco, CA, August 7-11, 1989.

[4] Pearl, J., *Probabilistic Reasoning in Intelligent Systems: Networks of Plausible Inference*, Morgan Kaufmann Publishers, Inc., 1988.

[5] Uemura, Y., and Sakawa, M., "A Decision rule on Possibility Distributions of Fuzzy Events," *Cybernetics and Systems*, vol.24, no.2, 1993.

[6] Yang, C. C., and Cheung, K. M., "Fuzzy Bayesian Analysis with Continuous-Valued Evidence," *Proceedings of IEEE International Conference on Systems, Man, and*

Cybernetics, Vancouver, British Columbia, Canada, October 22-25, 1995.

[7] Zadeh, L. A., "Fuzzy Sets," *Information and Control*, vol.8, no.3, June, 1965.

[8] Zadeh, L. A., "Fuzzy Algorithms," *Information and Control*, vol.12, no.2, February, 1968.

[9] Zadeh, L. A., "Quantization Fuzzy Semantics," *Information Sciences*, vol.3, no.2, April, 1971.