퍼지 사전 모수에 관한 베이지안 가설검정

Hypotheses testing of Bayes' theorem for fuzzy prior parameters

Man-Ki Kang[*], Gue-Tak Chio[**]
[*]Dept. Information Statistics, Dongeui University
(mkkang@dongeui.ac.kr)
[**] Kyungnam College of Information & Technology

Abstract

We have fuzzy hypotheses testing from Bayesian statistics with ideas from fuzzy sets theory to generalize Bayesian methods both for samples of fuzzy data and for prior distributions with non-precise parameters. Appling the principle of agreement index, the posterior odds ratio in the favor of hypotheses H_0 is equal to product of the fuzzy odds ratio and the fuzzy likelihood ratio. If the posterior odds ratio exceeds the grade judgement, we accept the hypothesis H_0 for the degree.

Key words: hypotheses testing, Bayesian statistics, prior distributions, posterior odds ratio

1. Preliminaries

We have fuzzy hypotheses testing from Bayesian statistics with ideas from fuzzy sets theory to generalize Bayesian methods both for samples of fuzzy data and for prior distributions with non-precise parameters. The posterior odds ratio in the favor of hypotheses H_0 is equal to product of the fuzzy odds ratio and the fuzzy likelihood ratio. If the posterior odds ratio exceeds the grade judgement, we accept the hypothesis H_0 for the ratio degree.

Consider a stochastic quantity $\tilde{x}$ which follows a distribution with probability density function(p.d.f.) belong to a parametric family:

$$\tilde{x} \sim f(x \mid \theta),\ \boldsymbol{\theta} = (\theta_1, \cdots, \theta_m) \in \Theta \subseteq R^m.$$

In Bayesian statistics one assume that the parameter $\boldsymbol{\theta}$ of the p.d.f. $f(x \mid \boldsymbol{\theta})$ is a stochastic quantity, too.

A prior $\boldsymbol{\theta}$ is a distributed according to some p.d.f. $\pi(\boldsymbol{\theta} \mid \eta_0)$. The prior parameter

$\eta_0 \in E \subseteq R^p$ has to be known.

Given sample $x=(x_1, \cdots, x_n)$ of n stochastically independent observation of $\tilde{x}$ the prior p.d.f. of θ is updated by means of Bayes' theorem:

$$\pi(\theta | \eta_0, x) \propto g_n(\theta | \eta_0, x),$$

$$g_n(\theta | \eta_0, x) = \pi(\theta | \eta_0) \cdot L(\theta | x)$$

$$= \pi(\theta | \eta_0) \cdot \prod_{I=0}^{n} f(x_i | \theta)$$

$\pi(\theta | \eta_0, x)$ is called the posterior p.d.f. of θ, $g_n(\theta | \eta_0, x)$ is called the non-normalized posterior p.d.f.

II. Bayes' theorem for fuzzy continuous parameter

Fuzziness may enter in two way: through fuzziness of data and through fuzziness of prior parameter.

We consider that Bayes' theorem for fuzzy data by introducing a fuzzy value posterior p.d.f. and extend this approach to deal with both sources of fuzziness.

Also we will use the notation, $\tilde{x} = (\tilde{x}_1, \cdots, \tilde{x}_n)$ will denote a fuzzy sample, $\tilde{\eta}_0$ a fuzzy parameter of prior, $\phi_{\tilde{x}}(x)$ and $\phi_{\tilde{\eta}_0}(\eta_0)$ will denote their characterizing or membership functions, $(C(\tilde{x})_\alpha, \alpha \in (0,1]$ and $(C(\tilde{\eta})_\alpha, \alpha \in (0,1])$ will denote the corresponding α-cut representations.

If the only the data are fuzzy, then only functional value of the likelihood function, denote by $L(\theta | \tilde{x})$, is fuzzy, where the functional value of the prior is a precise number. Therefore the α-contours of the non-normalized posterior $g_n(\theta | \eta_0, \tilde{x})$ are related to the α-contours $(L)_\alpha^L(\theta)$ and $(L)_\alpha^U(\theta)$ of the likelihood simple by

$$(g_n)_\alpha^L(\theta) = \pi(\theta | \eta_0) \cdot (L)_\alpha^L(\theta),$$

$$(g_n)_\alpha^U(\theta) = \pi(\theta | \eta_0) \cdot (L)_\alpha^U(\theta).$$

Where the contours of the fuzzy valued likelihood function are easy to determine or not.

If $\tilde{x}$ is a minimum rule fuzzy sample, then α-cuts of the single data points and the α-contours of the likelihood are given by

$$(L)_\alpha^L(\theta) = \min_{x \in C(\tilde{x})_\alpha} L(\theta | x) = \prod_{I=0}^{n} (f_i)_\alpha^L(\theta) \quad (2\text{-}1)$$

$$(L)_\alpha^U(\theta) = \max_{x \in C(\tilde{x})_\alpha} L(\theta | x) = \prod_{I=0}^{n} (f_i)_\alpha^U(\theta) \quad (2\text{-}2)$$

with

$$(f_i)_\alpha^L(\theta) = \min_{x \in C(\tilde{x})_\alpha} f(x | \theta),$$

$$(f_i)_\alpha^U(\theta) = \max_{x \in C(\tilde{x})_\alpha} f(x | \theta). \quad (2\text{-}3)$$

This result is of both practical and theoretical importance. the practical importance lies in the computational aspect that instead of minimizing and maximizing the likelihood function in the n-dimensional argument x, for a minimum rule fuzzy sample this n-dimensional optimization problem is deduced to the one dimensional problem of minimizing and maximizing the p.d.f. $f(x | \theta)$ over $x \in C(\tilde{x}_i)_\alpha$. This important property

For example, we assume that the stochastic quantity $\tilde{x}$ follows an exponential distribution:

$$f(x | \theta) = \theta \exp(-\theta x) \quad (2\text{-}4)$$

with unknown parameter $\theta \in R^+$. an conjugate prior for θ is given by a gamma density with parameter $\eta_0 = (v_0, \beta_0)$:

$$\pi(\theta | v_0, \beta_0) = \frac{\beta_0^{v_0}}{\Gamma(v_0)} \theta^{v_0 - 1} \exp(-\beta_0 \theta) \quad (2\text{-}5)$$

Assume that n fuzzy observation $\tilde{x}_i$ are available and they are combined by the minimum rule to a fuzzy sample $\tilde{x}$.

To obtained the α-contours of the non-normalized posterior we have to minimize and maximize

$f(x|\theta)$ for fixed θ over

$$x \in C(\tilde{x}_i) = [C_L(\tilde{x}_i)_\alpha, C_U(\tilde{x}_i)_\alpha] \quad (2\text{-}6)$$

for $i=1,\cdots,n$, as $f(x|\theta)$ is decreasing function in the x the argument leading to the mimimum and to the maximum, respectively, simple are given by

$C_U(\tilde{x}_i)_\alpha$ and $C_L(\tilde{x}_i)_\alpha$. Therefore we obtain the following α-contours of the fuzzy valued non-normalized posterior $g_n(\theta|\ \eta_0, \tilde{x})$:

$$(g_n)^L_\alpha(\theta) = \frac{\beta_0^{v_0}}{\Gamma(v_0)}\ \theta^{v_0+n-1}\ \cdot,$$

$$\exp(-(\beta_0 + \sum_{i=1}^{n} C_U(\tilde{x}_i)_\alpha)\theta) \quad (2\text{-}7)$$

$$(g_n)^U_\alpha(\theta) = \frac{\beta_0^{v_0}}{\Gamma(v_0)}\ \theta^{v_0+n-1}\ \cdot.$$

$$\exp(-(\beta_0 + \sum_{i=1}^{n} C_L(\tilde{x}_i)_\alpha)\theta) \quad (2\text{-}8)$$

III. Hypotheses testing

We consider the case of testing a simple hypotheses

$$H_0 : \theta = \theta_0$$

against a alternative

$$H_1 : \theta = \theta_1$$

where θ_0 and θ_1 preassigned constants.

W assume that H_0 and H_1 are mutually exclusive and exhaustive hypotheses.

Let $\tilde{x}_1 = \tilde{x}_1(X_1, \cdots, X_N)$ and $\tilde{x}_2 = \tilde{x}_2(X_1, \cdots, X_N)$ denote appropriate test statistics based upon two samples of N observations, respectively.

By Bayes' theorem, the posterior probability of H_0, given the observation data $\tilde{x}_1$ and for hypothesis H_1 from $\tilde{x}_2$, we have

$$\frac{\pi(\theta_0|\ \eta_0, x_1)}{\pi(\theta_1|\ \eta_0, x_2)} = \frac{\pi(\theta_0|\ \eta_0)\cdot L(\theta|x_1)}{\pi(\theta_1|\ \eta_0)\cdot L(\theta|x_2)} \quad (3\text{-}1)$$

That is, the posterior odds ratio in favor of H_0 is equal to the product of the prior odds and the likelihood ratio.

If the posterior odds ratio exceeds unity, we accept H_0; otherwise, we reject H_0 in favor of H_1.

IV. Example

Two fuzzy samples were derived by taking the simulated value fuzzy number $\tilde{x}(l_i, r_i)$. The parameter l_i and r_i determining the amount of fuzziness for the first fuzzy sample are samller than the corresponding parameters for the fuzzy sample.

A Bayesian analysis was carried out with the prior's parameters equal to $v_0 = 2, \beta_0 = 1$. For the fuzzy sample the fuzziness of the data has a considerable impact on the non-normalized posterior.

If we have random sample from population 1 as:

l	r
0.2	0.5
1.31	1.61
1.65	1.85
0.77	0.97
0.66	0.91
2.3	2.65
2.23	2.43
1.51	1.76
0.59	0.84
0.38	0.78

Also, we have another random sample from population 2 as:

l	r
0.2	0.9
0.91	1.71
1.35	2.3
0.52	1.27
0.46	1.21
1.9	2.85
2.03	2.73
1.31	1.96
0.19	1.14
0.18	0.98

From random sample 1, we have fuzzy l, r fuzzy number as;

$$g_n(\theta \mid \eta_0, \tilde{x}_1) = (3.432E-09, 8.121E-10, 1.396E-10) \quad (3\text{-}2)$$

From random sample 2, we have l, r fuzzy number as;

$$g_n(\theta \mid \eta_0, \tilde{x}_2) = (3.581E-08, 8.833E-10, 1.713E-11) \quad (3\text{-}3)$$

Thus we have fuzzy number posterior odds ratio as;

$$(0.005, 1.040, 200.336) \quad (3\text{-}4)$$

Because of division process of two fuzzy number, the right number of the fuzzy number is divergence in the province of greater than 1. We will test the model with the modified fuzzy number as;

$$(0.005, 1.040, 2.075) \quad (3\text{-}5)$$

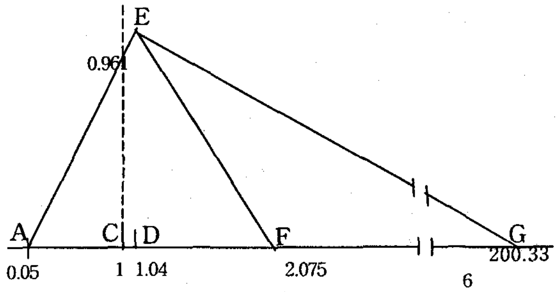
We accept hypothesis H_0, because the value of center 1.04. is great than 1 in non fussy case.

Also, for the fussy case, we accept the H_0 when the area is greater than 1 and reject the case when the area is less than 1.

Thus we accept the hypothesis H_0 for the fuzzy posterior odds ratio degree

$$1 - \frac{area(triangularABC)}{area(trinagularAEF)} = 0.5392 \quad (3\text{-}6)$$

as seen in Figure.



[Figure]

References

[1] S. Fruhwirth-Schnatter, On Fuzzy Bayesia n Inference, *Fuzzy Sets and Systems* 60(1993), 41-58.

[2] P. X. Gizegorzewski, Testing Hypotheses with vague data, *Fuzzy Sets and Systems*. 112 (2000), pp.501-510

[3] M. K. Kang, G. T. Choi and C. E. Lee, On Statistical for Fuzzy Hypotheses with Fuzzy Data, *Proceeding of Korea Fuzzy Logic and Intelligent System Society Fall Conference*, Vol. 10, Num. 2, (2000)

[4] M. K. Kang, G. T. Choi and S. I .Han, A Bayesian Fuzzy Hypotheses testing with Loss Function, *Proceeding of Korea Fuzzy Logic and Intelligent Systems Society Fall Conference,* Vol 13, Num. 2, (2002)

[5] F. Li, C. Wu, J. Qiu and L. Su, Platform type fuzzy number a separability of fuzzy number space, *Fuzzy Sets and systems* 117 (2000), 347-353.

[6] N. Watanabe, T. Imaizumi, A Fuzzy Statistical Test of Fuzzy Hypotheses, *Fuzzy Sets and systems* 53 (1993), pp.167-178.

[7] S. J. Press, *Bayesian Statistics: Principle, model and Application,* John Willy & Sons, 1999.