Fuzziness and Bayesian Analysis in Engineering

Matthias Stein[1], Michael Beer[2,4], and Vladik Kreinovich[3]
[1] Bechtel OG&C Offshore, Houston, USA
[2] Institut for Risk & Uncertainty, University of Liverpool, Liverpool, UK
[3] Department of Computer Science, University of Texas at El Paso, El Paso, TX, USA
[4] Corresponding author, e-mail: mbeer@liverpool.ac.uk

Abstract

An engineering analysis requires a realistic quantification of all input information. The amount and quality of the available information dictates the uncertainty model and its associated quantification concept. For inconsistent information, a distinction between probabilistic and non-probabilistic characteristics is beneficial. In this distinction, uncertainty refers to probabilistic characteristics and non-probabilistic characteristics are summarized as imprecision. When uncertainty and imprecision occur simultaneously, the uncertainty model fuzzy randomness appears useful. In a Bayesian approach the fuzzy probabilistic model provides the opportunity to take account of imprecision in data and in prior expert knowledge. The Bayesian approach ex-tended to inconsistent information is demonstrated by means of an example.

Keywords: fuzzy-Bayes, fuzzy random variables, imprecise data, imprecise probabilities, uncertainty quantification

1. Introduction

Whenever decision making in engineering is based on results from a numerical analysis, it is essential that the available information is reflected properly in these results. This requires a realistic modeling of the available information without distorting it or ignoring certain aspects. In engineering practice, the information basis usually consists of plans, drawings, measurements, observations, experiences, expert knowledge, codes and standards, and so on. Hence, it is frequently not certain or precise but rather imprecise, diffuse, fluctuating, incomplete, fragmentary, vague, ambiguous, dubious, or linguistic. In addition, data and expert assessments may not be in full agreement. This type of inconsistent information requires a generalized uncertainty model to minimize the risk from modeling errors under uncertainty in engineering computations and prognoses. Shortcomings, in this regard, may lead to biased computational results with an unrealistic accuracy and may therefore lead to wrong decisions with the potential for associated serious consequences. Uncertainty modeling has thus already become an engineering task of great importance and interest.

In the present study, fuzzy probability theory (Beer, 2009a) is selected for the simultaneous treatment of uncertainty and imprecision. The fuzzy probabilistic model provides a high degree of generality and flexibility, as it combines probabilistic modeling with fuzzy modeling. It enables to take into account an entire set of plausible probabilistic models complying with the underlying information. This is mathematically realized with the aid of fuzzy variables for the distribution parameters and for the description of the distribution type. The associated fuzzy probability distribution functions $\tilde{F}_{\tilde{X}}(x)$ represent a bunch of plausible traditional distribution functions with fuzzy parameters as bunch parameters; their functional values are fuzzy numbers.

In order to utilize the features and benefits of generalized uncertainty models, which include both uncertainty and imprecision, suitable quantification methods are needed to capture the available information in each particular case in the most realistic manner. This can be achieved by merging methods of traditional mathematical statistics and Bayesian statistics with statistics for imprecise data and set-theoretical methods. So far, the combination of Bayesian methods with imprecise information has not been investigated extensively. Developments in this direction, summarized with the term fuzzy Bayes methods, have been published in (Viertl & Hareter 2004, Viertl 2008, Viertl 2011). This approach involves averaging elements regarding the propagation of the imprecision during the update using the bounds of the imprecision. In the present study we pursue an alternative, numerical approach to reflect the imprecision of the input information in the predicted distribution using concepts for propagation of imprecision in compliance with the fuzzy probabilistic model in (Beer 2009a).

2. Bayesian Update with Fuzzy Information

2.1 Quantification options

We start from the Bayesian theorem

$$f_{\underline{\Theta}|\underline{X}=\underline{x}}\left(\underline{\theta}|\underline{x}\right)=\frac{f_{\underline{X}|\underline{\Theta}=\underline{\theta}}\left(\underline{x}|\underline{\theta}\right)\cdot g_{\underline{\Theta}}\left(\underline{\theta}\right)}{f_{\underline{X}}\left(\underline{x}\right)} \tag{1}$$

and identify four options to develop a fuzzy Bayesian update.

The first option covers the case where no component of Eq. (1) is affected by fuzziness, but where fuzziness is induced by an interval estimation of the posterior parameters. The subjective selection of the estimator, the subjective selection of the confidence level for interval estimation, and the selection of the type of the interval are sources of fuzziness. The parameter estimation can be performed for a variety of selections, and the results are merged to a fuzzy parameter $\underline{\tilde{\theta}}$ eventually leading to a fuzzy random variable $\underline{\tilde{X}}$.

The second option covers the case of an imprecise likelihood function, where according to prior information the type of the distribution of $\underline{X}$ and the parameters other than $\underline{\theta}$ cannot be specified precisely. The random variable $\underline{X}$ is then already present as a fuzzy random variable $\underline{\tilde{X}}$ and enters Eq. (1). This yields, again, a fuzzy random variable $\underline{\tilde{X}}$ after the update.

The third option covers the case where fuzziness is induced due to imprecision in the prior distribution.

Option four is concerned with the scenario of a Bayesian update with imprecise data $\underline{\tilde{x}}$.

The four basic options can be combined with each other to build further quantification options in the gaps between the basic options. Herein we focus on the latter two options and their combination.

2.2 Imprecise prior distribution

The likelihood function $f_{\underline{X}|\underline{\Theta}=\underline{\theta}}\left(\underline{x}|\underline{\theta}\right)$ is known precisely. Only vague information is, however, available for the specification of the prior distribution so that only bounding functions for $g_{\underline{\Theta}}\left(\underline{\theta}\right)$ can be formulated. This prompts the consideration of the set of all plausible prior distributions. The parameters of $g_{\underline{\Theta}}\left(\underline{\theta}\right)$ may therefore be specified with the aid of “subjectively assessed intervals” to represent fuzzy sets. The prior distribution then appears as fuzzy probability distribution $\tilde{g}_{\underline{\tilde{\Theta}}}\left(\underline{\theta}\right)$ in Eq. (1). The result obtained from Eq. (1) is a fuzzy probability density function $\tilde{f}_{\underline{\tilde{\Theta}}|\underline{X}=\underline{x}}\left(\underline{\theta}|\underline{x}\right)$ as the basis to estimate the parameters $\underline{\theta}$ of the random variable $\underline{X}$. As the parameters of $\tilde{f}_{\underline{\tilde{\Theta}}|\underline{X}=\underline{x}}\left(\underline{\theta}|\underline{x}\right)$ are fuzzy variables, this estimation yields fuzzy values $\underline{\tilde{\theta}}$, which eventually lead to a fuzzy random variable $\underline{\tilde{X}}$. Fuzzy randomness is induced by the assumption of $\tilde{g}_{\underline{\tilde{\Theta}}}\left(\underline{\theta}\right)$.

2.3 Imprecise data

The distribution parameters of both the likelihood function $f_{\underline{X}|\underline{\Theta}=\underline{\theta}}\left(\underline{x}|\underline{\theta}\right)$ and the prior density function $g_{\underline{\Theta}}\left(\underline{\theta}\right)$ are known precisely. The sample for the Bayesian update is, however, comprised of imprecise data and is taken into consideration as a fuzzy sample $\left(\tilde{x}_1,...,\tilde{x}_n\right)$. In Eq. (1) the fuzzy sample elements $\tilde{x}_i$ are evaluated with the function $f_{\underline{X}|\underline{\Theta}=\underline{\theta}}\left(\underline{x}|\underline{\theta}\right)$ leading to fuzzy functional values $\tilde{f}_{\underline{X}|\underline{\Theta}=\underline{\theta}}\left(\tilde{\underline{x}}|\underline{\theta}\right)$. This yields a fuzzy probability density function $\tilde{f}_{\tilde{\underline{\Theta}}|\underline{X}=\underline{x}}\left(\underline{\theta}|\underline{x}\right)$, again, and hence a fuzzy random variable $\tilde{\underline{X}}$. Fuzzy randomness is induced via the fuzzy sample.

3. Illustrative Application

3.1 Problem description

A sample of size 20 is available for the compressive strength f_c of concrete. This sample is taken from (Beer 2009b), where it is presented (i) as precise data and (ii) as fuzzy data with a spread of ±2 N/mm^2 and then used for a non-Bayesian fuzzy probabilistic quantification. The respective results are used, herein, for the purpose of comparison.

The compressive strength of concrete f_c is modeled as a normal random variable X,

$$f_c = X \sim N\left(\mu_X, {\sigma_X}^2\right) \qquad (2)$$

Prior distributions are assumed for the distribution parameters μ_X and σ_X, which are then updated based on the data. With $\underline{\Theta}$ as general notation for uncertain distribution parameters and $g_{\underline{\Theta}}\left(\underline{\theta}\right)$ being their joint probability density function, the Bayesian estimates for μ_X and σ_X are obtained as expected values of $\underline{\Theta}$ based on the posterior distribution $f_{\underline{\Theta}|\underline{X}=\underline{x}}\left(\underline{\theta}|\underline{x}\right)$,

$$\mu_X = E\left[\Theta_1\right], \sigma_X = E\left[\Theta_2\right] \qquad (3)$$

We assume that the prior distribution for $\underline{\Theta}$ is given by the marginal pdf's (probability density functions) for Θ_1 and Θ_2, and that Θ_1 and Θ_2 are statistically independent of one another. Two cases will be considered:

(1) the case “normal”, where $\Theta_1 \sim N\left(\mu_{\Theta_1}, \sigma_{\Theta_1}\right)$ and $\Theta_2 \sim N\left(\mu_{\Theta_2}, \sigma_{\Theta_2}\right)$,

(2) the case “uniform”, where $\Theta_1 \sim U\left(a_{\Theta_1}, b_{\Theta_1}\right)$ and $\Theta_2 \sim U\left(a_{\Theta_2}, b_{\Theta_2}\right)$.

These cases are investigated with the aid of fuzzy variables for the numerical description of imprecision. This leads to fuzzy probability distribution functions for various random variables involved. The numerical processing of the fuzziness is realized with α-level optimization as described in (Möller and Beer 2004).

3.2 Imprecise prior distribution

Let several experts agree upon a most plausible value for the parameters and some intervals, which cover the range of their individual opinions. From this information, fuzzy triangular numbers can be constructed,

$$\begin{aligned}\tilde{\mu}_{\Theta_1} &= \langle 26.5, 28.5, 30.5\rangle\, N/mm^2,\quad \tilde{\sigma}_{\Theta_1} = \langle 1.0, 2.0, 3.0\rangle\, N/mm^2 \\ \tilde{\mu}_{\Theta_2} &= \langle 3.5, 4.5, 5.5\rangle\, N/mm^2,\quad \tilde{\sigma}_{\Theta_2} = \langle 0.75, 1.0, 1.5\rangle\, N/mm^2\end{aligned} \qquad (4a)$$

for the "normal" case as and

$$\begin{aligned}\tilde{a}_{\Theta_1} &= \langle 20.0, 27.0, 28.0\rangle\, N/mm^2,\quad \tilde{b}_{\Theta_1} = \langle 29.0, 30.0, 36.0\rangle\, N/mm^2 \\ \tilde{a}_{\Theta_2} &= \langle 2.0, 3.0, 4.0\rangle\, N/mm^2,\quad \tilde{b}_{\Theta_2} = \langle 5.0, 6.0, 7.0\rangle\, N/mm^2\end{aligned} \qquad (4b)$$

for the "uniform" case.

For each case, the associated fuzzy probability distribution $\tilde{g}_{\tilde{\Theta}}(\underline{\theta})$ is then used for the Bayesian update. The result is a fuzzy probability density function $\tilde{f}_{\tilde{\Theta}|\underline{X}=\underline{x}}(\underline{\theta}|\underline{x})$. As the parameters of $\tilde{f}_{\tilde{\Theta}|\underline{X}=\underline{x}}(\underline{\theta}|\underline{x})$ are fuzzy variables, Eq. (3) yields fuzzy values $\tilde{\mu}_X$ and $\tilde{\sigma}_X$, which eventually lead to a fuzzy random variable $\underline{\tilde{X}}$ for the compressive strength. Consequently, the 5% quantile is obtained as a fuzzy variable. Figure 3 illustrates the Bayesian update for the case of a "normal" fuzzy prior distribution and precise data. To investigate the effect of the sample size, the quantification of $\tilde{\mu}_X$ and $\tilde{\sigma}_X$ was again carried out for a sample size of n = 7 and n = 20. It can be observed that the fuzziness, which is induced solely by the assumption of $\tilde{g}_{\tilde{\Theta}}(\underline{\theta})$, is significantly less for n = 20. This corresponds well with the standard Bayesian approach, in which the influence of the prior distribution decays with increasing sample size. The fuzzy parameters $\tilde{\mu}_X$ and $\tilde{\sigma}_X$ show some interactive dependency.

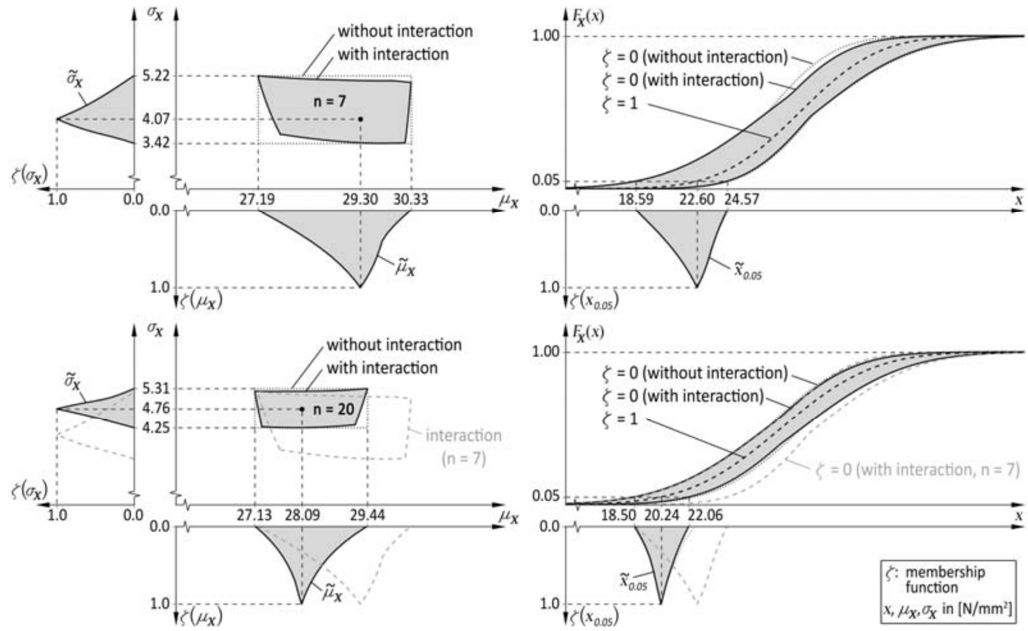

Figure 3. Bayesian update for "normal" fuzzy prior; interaction between fuzzy distribution parameters; resulting fuzzy quantiles.

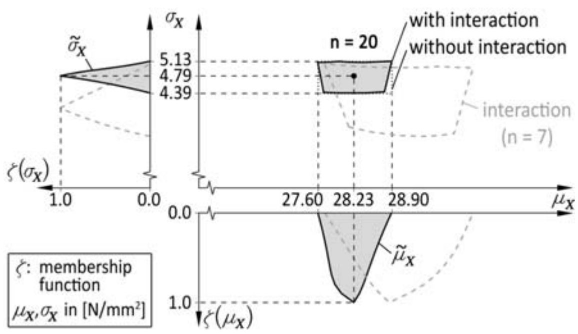

Figure 4. Case "uniform" fuzzy prior.

Figure 4 depicts the fuzzy parameters $\tilde{\mu}_X$ and $\tilde{\sigma}_X$ and their interaction for the case of a "uniform" fuzzy prior. Again, the fuzziness decays visibly with increasing sample size. A significant difference between the case of a "normal" prior distribution and a "uniform" prior distribution can be observed in the shape of the membership functions for $\tilde{\mu}_X$ and $\tilde{\sigma}_X$. The concave shape of the membership functions in the case of a “normal” prior distribution implies that a small reduction of input imprecision in $\tilde{g}_{\tilde{\Theta}}(\underline{\theta})$ results in a strong reduction of imprecision in the estimation results. In contrast to this, the convex shape of the membership functions in the case of a “uniform” prior distribution implies that a reduction of imprecision in $\tilde{g}_{\tilde{\Theta}}(\underline{\theta})$ is less effective to reduce imprecision in the estimation results. For both cases it can be seen that the exact result is fully covered if the interaction is neglected.

3.3 Imprecise data

Suppose that a fuzzy sample with elements $\tilde{x}_i$ is available according to (Beer 2009b). The Bayesian update then leads to the fuzzy probability density function $\tilde{f}_{\tilde{\Theta}|\underline{X}=x}\left(\underline{\theta}|\underline{x}\right)$, and consequently again to fuzzy values $\tilde{\mu}_X$ and $\tilde{\sigma}_X$ as the parameters for the fuzzy random variable $\underline{\tilde{X}}$. The prior distribution is specified precisely with distribution parameters according to the middle numbers in Eqs. (4a) and (4b).

Figure 5 illustrates the Bayesian update with a "normal" prior distribution and the imprecise data for n = 7 and n = 20. It can be observed that the fuzziness in the estimation results, which is now exclusively induced by the sample elements $\tilde{x}_i$, increases with increasing sample size. Although non-intuitive, this again corresponds well with the standard Bayesian approach, in which the influence of the now precise prior distribution decays with increasing sample size. As the estimation results converge towards the statistical solution (data only), the imprecision in the estimation results as well converges towards the solution from statistics with imprecise data (see Beer 2009b).

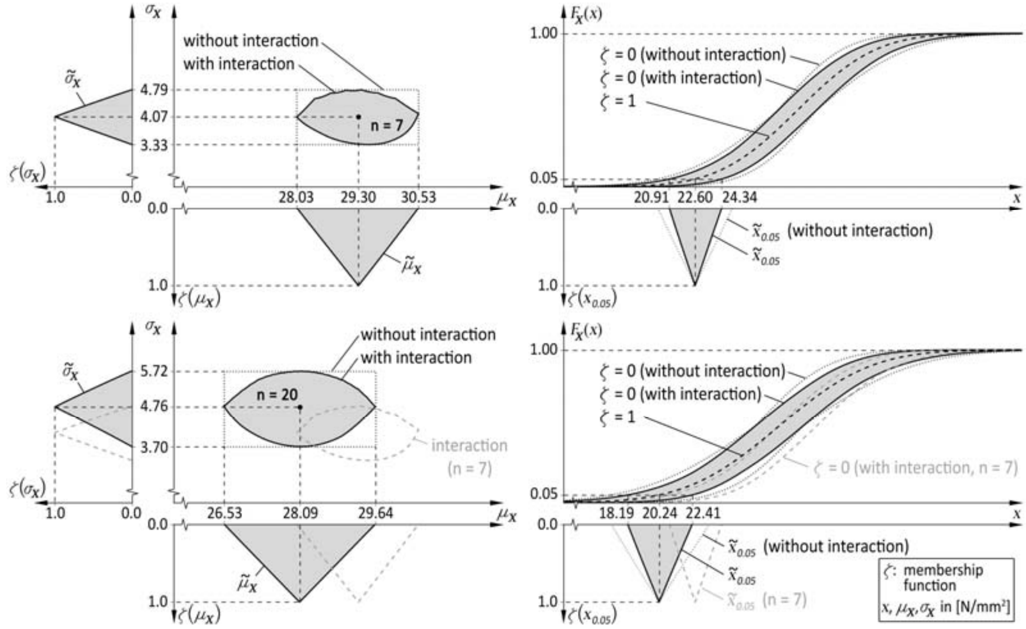

Figure 5. Bayesian update for "normal" prior and imprecise data; interaction between fuzzy distribution parameters; fuzzy quantiles.

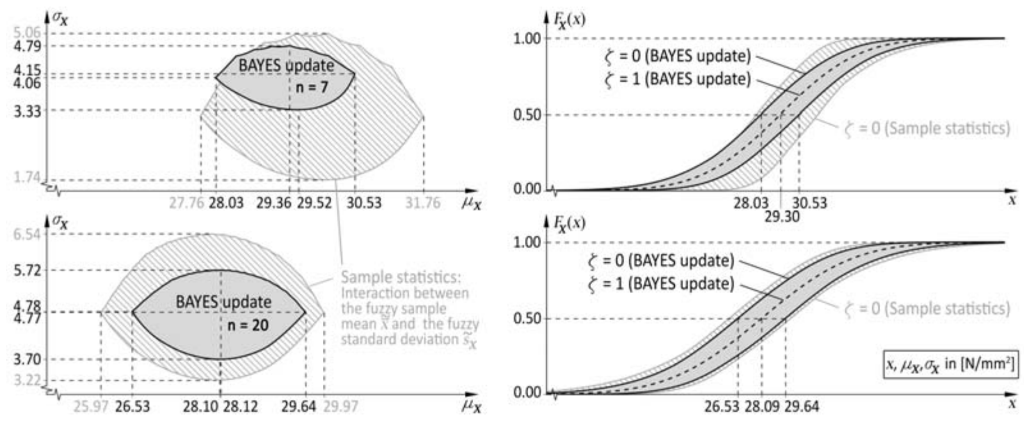

Figure 6. Comparison of the estimation results from Bayesian update with imprecise data with the results from sample statistics with imprecise data.

Figure 6 shows that comparison of the estimation results from Bayesian update with imprecise data with the results from statistics with imprecise data. It can be observed that the use of precise prior information, if available, is significantly reducing the imprecision of the results. This is especially the case if only rare and imprecise data are available, i.e. in the example for the case of n = 7. The effect decays with increasing sample size. The same effects appear in the case “uniform”. Ignoring the interaction between the fuzzy parameters $\tilde{\mu}_X$ and $\tilde{\sigma}_X$ increases significantly the imprecision in the quantiles. This effect is less severe for the "uniform" prior distribution.

3.4 Imprecise prior distribution and imprecise data

The two basic options for the Bayesian update with imprecise information as presented in Sections 3.2 and 3.3 can be combined. In such combination the effects of the imprecision in the prior distribution and of the imprecision in the data on the estimation results counterbalance one another. Consequently, the imprecision in the estimated fuzzy parameters $\tilde{\mu}_X$ and $\tilde{\sigma}_X$ does not significantly change with the sample size. Whilst for small sample size the effect of the imprecision in the prior distribution is dominant, the impact of imprecision in the data is dominant for larger sample sizes.

4. Conclusions

Inconsistent information represents a common problem in engineering practice. This information is not appropriate for a plain evaluation by means of traditional statistics. A proper evaluation and a suitable numerical description are, however, required to obtain realistic results in a structural analysis, which are often the basis for engineering decision. To achieve this goal, the model fuzzy randomness is utilized, which enables a separate and simultaneous treatment of statistical uncertainty and imprecision. Due to the variety of possible forms of available information, a general quantification algorithm cannot be formulated. In the case of some subjective probabilistic information next to measured data, the quantification can be realized with a Bayesian update. If either the data or the prior probabilistic information inherits imprecision, the Bayesian update can be extended to fuzzy Bayes methods to quantify a fuzzy random variable based on all available information. It has been shown, that the Bayesian update with imprecise information retains the attractive properties of the standard Bayesian approach. The imprecision of the input information is reflected as imprecision in the estimation results without being averaged out. Hence, the predicted distribution exhibits both uncertainty and imprecision.

References

Beer, M., 2009a. Fuzzy Probability Theory In: Meyers, R. (ed.), Encyclopedia of Complexity and Systems Science, Vol 6, 4047–4059, New York: Springer.

Beer, M., 2009b. Engineering quantification of inconsistent information, *Int. J. Reliability and Safety*, 3(1/2/3): 174–197.

Möller, B. & Beer, M., 2004. Fuzzy Randomness – Uncertainty in Civil Engineering and Computational Mechanics. Springer, Berlin.

Viertl, R. & Hareter, D., 2004. Generalized Bayes-theorem for non-precise a-priori distribution, *Metrika*, 59 (3): 263–273.

Viertl, R., 2008. Foundations of Fuzzy Bayesian Inference. *Journal of Uncertain Systems*, 2 (3):187-191.

Viertl, R., 2011. Statistical Methods for Fuzzy Data. Wiley, Chichester.