FUZZY PROPERTIES OF REAL RANDOM VARIABLES

Miroslav Pokorný, Vratislava Mošová, Zdeňka Krišová, Vladimíra Mádrová Miroslav Pokorný, Department of Informatics and Applied Mathematics, Moravian University College Olomouc, Tř. Kosmonautů 1, 779 00 Olomouc, Czech Republic, e-mail: miroslav.pokorny@mvso.cz Vratislava Mošová, Department of Informatics and Applied Mathematics, Moravian University College Olomouc, Tř. Kosmonautů 1, 779 00 Olomouc, Czech Republic, e-mail: Vratislava.mosova@mvso.cz Zdeňka Krišová, Department of Informatics and Applied Mathematics, Moravian University College Olomouc, Tř. Kosmonautů 1, 779 00 Olomouc, Czech Republic, e-mail: Zdenka.krisova@mvso.cz Vladimíra Mádrová, Department of Informatics and Applied Mathematics, Moravian University College Olomouc, Tř. Kosmonautů 1, 779 00 Olomouc, Czech Republic, e-mail: Zdenka.krisova@mvso.cz

Abstract:

In practice, we need to consider fuzzy variables affected by stochastic factors. They cannot be formalized either using solely fuzzy or stochastic approaches. In a number of analyses of practical systems, it is therefore necessary to use an integrated approach – fuzzy-stochastic - and formalize the fuzzy random variables. A fuzzy random variable can be regarded as a random variable measured under uncertain conditions, i.e. was not obtained under strictly defined experimental conditions. To estimate such fuzzy random variables statistical methods can be applied, they are, however, extended by including the fuzzitivity of random data. The paper addresses the issue of the theory of fuzzy random variables, determining their functional and numerical fuzzy characteristics and presents a numerical example.

Key words:

stochasticity, fuzzitivity, fuzzy set, fuzzy-stochastic value, functional characteristics, numerical characteristics

JEL: C51, C65

1 Introduction: Heading for the introductory chapter

Information and results from statistical analyses can be, to a significant degree, devalued by the influence of non-negligible vague certainty of the input variables and calculation models. Therefore it needs to be decided, from which moment in the statistical analysis the necessity arises to recognize and quantify also the influence of the subjective (vague, fuzzy) uncertainty.

It is possible to state that the existence of fuzzy stochasticity can be tangible in practical cases, if:

- The extent of the selection sets is small and lacks further a priori information on the static properties of the measured variable,
- Statistical data has the property of fuzzitivity, i.e. has uncertain accuracy
- Statistical data was gained under uncertain, non-defined or non-reproduced circumstances.

In order to express the degree of vagueness of a random variable, the following non-parametrical tests of statistical hypotheses of the properties and parameters of the selection set can be used:

- Set randomness test zero hypothesis H0: "Elements of the set are random"
- Set homogeneity test test of a set including the elements (a d) and zero hypothesis
 H0: "The selection set consists of two subsets (a b) and (c d), originating from an identical distribution"
- Kolmogorov-Smirnov set distribution type test the type of the set distribution function is estimated using the Kolmogorov-Smirnov nonparametric test of equality. If the type of the set distribution cannot be unambiguously estimated, the statistical prerequisite of identical and independent distribution of the selection set is not fulfilled.

The statistical hypotheses tests enable verification or rejection of the premise of the stochastic character of the selection set.

If the selection set does not sufficiently prove its stochastic properties, a fuzzy stochastic approach has to be used in order to analyse it (Möller and Beer, 2004).

2 Methods

2.1 Fuzzy Random Variable

Fuzzy random variable is represented by random data being the bearer of complementary uncertainty - fuzzitivity (Möller and Beer, 2004), (Möller, 2004).

2.1.1 Fuzzy random variable model

Let us consider a space Ω of random phenomena ω (observation). Let us denominate by a fuzzy realisation the one-dimensional fuzzy random variable \widetilde{X} as $\widetilde{x}(\omega)$, $\omega \in \Omega$. Every fuzzy number \widetilde{x} is defined as a convex normal fuzzy set (Möller and Beer, 2004)

$$\widetilde{x} = \left\{ x; \mu_{\widetilde{x}} \left(x \right) \mid x \in X \right\}$$
(1)

where the membership function $\mu_{\tilde{x}}(x)$ is the function of pertinence of the fuzzy number \tilde{x} at

least partially continuous. The fuzzy random variable $\,\widetilde{\!X}\,$ is then defined as the fuzzy result of uncertain mapping

$$\widetilde{X}: \ \Omega \to F(\mathbb{R}^n) \tag{2}$$

where $F(R^n)$ is the set of all (normal) fuzzy numbers in R^n .



Fig. 1: Realisation of the one-dimensional fuzzy random variable X

Source: (Möller and Beer, 2004) (adapter)

Every realisation $x_i = x(\omega_i)$ of the real random variable *X* is in Fig.1 represented by black dots. The realisation is meanwhile one of the values of the bearer of the corresponding fuzzy number $x(\omega_i) \in \widetilde{x}(\omega_i)$ as a realisation of the fuzzy random variable \widetilde{X} .

The membership function of the triangular approximation of the fuzzy number \widetilde{A} with the meaning "probably α " is shown in Fug. 2. The numerical interval $\langle \alpha - c, \alpha + c \rangle$ denominates an uncertain area, where the values α with the degree of uncertainty $\mu_{\widetilde{A}}(x) \in \langle 0,1 \rangle$ occur (Novák, Perfilieva and Dvořák, 2016).

Fig. 2: Membership function of fuzzy number \widetilde{A}



The membership function is approximated by angular lines. The boundaries – bearer and the single elemented core of the fuzzy set – are given by the parameters

$$\widetilde{A}_{x} \approx \left[\alpha - c, \, \alpha, \, \alpha + c\right]. \tag{3}$$

2.2 Specifics of the Characteristics of Fuzzy Random Variable

2.2.1 Definition of numerical parameters of fuzzy random variable

The type of density of a probability distribution and the parameters of fuzzy random variable must be determined on the basis of an analysis of the selection set of fuzzy random variable \underline{X} . The following text discusses a one-dimensional fuzzy random variable \overline{X} . Let us mention the

relations for the parameters (moments) of its function of density of a probability distribution (Möller and Beer, 2004), (Möller, 2004).

The general moment of one-dimensional fuzzy random variable of the k-th order is defined as

$$\widetilde{m}_{k,x} = E\widetilde{X}^{k} = \int_{x=-\infty}^{x=+\infty} x^{k} \cdot \widetilde{f}(x) dx$$
(4)

For k = 1, we gain the relation for fuzzy medium value of the fuzzy random variable \widetilde{X} as

$$\widetilde{m}_{x} = E\widetilde{X} = \int_{x=-\infty}^{x=+\infty} \widetilde{f}(x)dx$$
(5)

the central moment of the one-dimensional fuzzy random variable of the k-th order is defined as

$$\widetilde{\zeta}_{k,x} = E(\widetilde{X} - \widetilde{m}_x)^k = \int_{x = -\infty}^{x = +\infty} (x - \widetilde{m}_x)^k \cdot \widetilde{f}(x) dx$$
(6)

if k = 2, then the relation of fuzzy sparsity for fuzzy random variable \widetilde{X} follows

$$\widetilde{\zeta}_{2,x} = D^2 \widetilde{X} = \int_{x=-\infty}^{x=+\infty} (x - \widetilde{m}_x)^2 \cdot \widetilde{f}(x) dx$$
(7)

and the fuzzy conclusive deviation of the random variable \widetilde{X} is given by the relation

$$\widetilde{\sigma}_{x} = \sqrt{D^{2}\widetilde{X}} = \sqrt{\int_{x=-\infty}^{x=+\infty} (x - \widetilde{m}_{x})^{2} \cdot \widetilde{f}(x) dx} .$$
(8)

The selection set of observed values is the source for the specification of pertinence. The base for the construction of the membership function is then a histogram or confidence interval. Subjective aspects can be used to correct its shape.

2.2.2 Analytical fuzzification of numerical parameters of the random variable

Confidence interval method

Presuming the existence of an estimate of the type of probability distribution of the random variable reduces the problem to the determination of numerical parameters for division. The parameters are modelled by fuzzy numbers. The medium value is defined by a point estimate and the boundaries of the fuzzy triangular fuzzy number are calculated as boundaries of the probability interval of mean value at the pertinence level α (Fig. 4), (Möller and Beer, 2004).

The estimates of the fuzzified mean value \tilde{m}_x and the estimate of the fuzzified sparsity \tilde{s}_x^2 are determined in the form of fuzzy numbers. The membership functions are approximated using angular lines – Fig. 3. The bounding points – the bearer and core of the fuzzy sets – are given by parameters

$$\widetilde{m}_{x} \approx \left[\Delta L, \overline{m}_{x}, \Delta R\right] \tag{9}$$

$$\widetilde{s}_{x}^{2} \approx \left[\Delta L, \overline{s}_{x}^{2}, \Delta R \right]$$
(10)

The values of cores \overline{m}_x and \overline{s}_x^2 are determined as the mean value and dispersion via conventional relations for the random variable with Gauss distribution

EMI, Vol. 10, Issue 1, 2018 ISSN: 1804-1299 (Print), 1805-353X (Online) www.emijournal.cz

$$\overline{m}_x = \frac{1}{n} \sum_{i=1}^n x_i \tag{11}$$

$$\bar{s}_x^2 = \frac{1}{n-1} \sum_{i=1}^n (x_i - \bar{m}_x)^2$$
(12)

the values of the left and right boundary of the bearer of both fuzzy numbers are determined as left and right boundary of the confidence interval of the mean value estimate and random variable sparsity with Gauss distribution on the level of significance $\alpha = 0.01$ (probability $P = (1 - \alpha) = 0.99$.

The membership functions $\mu(m_x)$ and $\mu(s_x^2)$ of the fuzzy sets \tilde{m}_x and \tilde{s}_x^2 are approximated using angular lines. The breaking points are determined by the left and right boundaries of confidence intervals used as the left and right boundaries of the α - sections. Fig. 3 shows the sections form the fuzzy set \tilde{m}_x .



Fig. 3: Approximated membership function m_x

Histogram method

The observed values x_i are divided into subsets X_k and a histogram is assembled. The histogram is then approximated by a suitable function.

Let us consider an approximation by linear function, leading to the gain of membership function in the shape of a triangle or a trapeze. The left and right boundary of the function bearer is gained as the intersection of the left and right approximation lines of the histogram with the *x*-axis. The intersection of the left and right approximation line determines the tip of the triangular approximation. In the case of the trapeze approximation, the left and right approximation lines determine the core of the fuzzy set (Fig. 4).

Fig. 4: Linear transformation of histogram



Source: own creation

The equations of left and right approximation lines are gained using the method of the smallest squares, minimalizing the difference of squares between the amount of observation in the *k*-th bar $n(X_k)$ and the functional value of approximation $\mu_A(x_{k,m})$, where $x_{k,m}$ is the mid point of the *k*-th bar (Möller, 2004)

$$\sum_{k} \left[n(X_k) - \mu_A(x_{k,m}) \right]^2 \to \min$$
(13)

The Y coordinate of the intersection of the left and right approximation line is normed to the value of 1.

$$\mu(x_i) = \left[\frac{p(x_i)}{\sup(p(x_i))}\right]^a, \ a \in (-\infty, \infty)$$
(14)

For a = 1, the function shape is purely standardised. For a > 1, the membership function narrows down, for a < 1, it widens (Fig. 5).

Fig. 5: Parametrization of membership function



(Möller and Beer, 2004) (adapted)

The triangular membership function can be at level 1 subsequently modified by core towards a trapeze function. It can also be modified by other conditions (subjective, expert).

3 Results

3.1 Numerical examples

3.1.1 Artificial Data Set Analysis

The efficiency of the proposed method was verified on the fuzzy-stochastic analysis of the randomly generated set N [3,1], containing 15 items. Such a set stems from a random variable with known (normal) distribution and small scale (Tab. I.), (Siegel, 2012).

Tab. I: Random file

| [2.68053, 2.15374, 3.33467, 1.37664, 2.74401, 2.40678, 4.52252, | |
|--|---|
| 2.47141, 4.12337, 2.28214, 4.59856, 4.34608 1.78819 1.63103 2.96888] | |
| Source: own creation | Ī |

The selection statistical parameters of the set $m_x = 2,89$ and $\sigma_x^2 = 1,13$. The shape of this histogram is shown in Fig. 6.



Fig. 6: Histogram of random file

Source: own creation

The shapes of the membership functions of the fuzzified statistical parameters \tilde{m}_x and $\tilde{\sigma}_x^2$, in the form of fuzzy numbers, were determined via the method of confidence intervals of the random variable with normal distribution and their allocation to the α -cut of the membership functions. The allocation of the confidence intervals and boundaries of the α -cuts is stated in Tab. II. (for \tilde{m}_x) and Tab. III. (for \tilde{s}_x^2), (Statistica, 2014).

| Tab. II: | Membe | rship | function | \widetilde{m}_{x} |
|----------|-------|-------|----------|---------------------|
|----------|-------|-------|----------|---------------------|

| Confidence level α | Membership function α-cut level | L – boundary | R – boundary |
|-----------------------|------------------------------------|--------------|--------------|
| 0,99 | 0,00 | 2,08 | 3,71 |
| 0,90 | 0,25 | 2,41 | 3,38 |
| 0,75 | 0,50 | 2,56 | 3,22 |
| 0,50 | 0,75 | 2,07 | 3,06 |
| | 1,00 | 2 | 89 |

Source: own creation

| Confidence level a | Membership function α-cut level | L- boundary | R-boundary |
|-----------------------|------------------------------------|-------------|------------|
| 0,99 | 0,00 | 0,51 | 3,90 |
| 0,90 | 0,25 | 0,67 | 2,42 |
| 0,75 | 0,50 | 0,79 | 1,92 |
| 0,50 | 0,75 | 0,93 | 1,56 |
| | 1,00 | 1, | 13 |

Tab. III: Membership function \tilde{s}_x^2

Source: own creation

Approximation of the membership functions \tilde{m}_x and $\tilde{\sigma}_x^2$ by the angular lines is stated in Fig.7. The non-symmetries of the membership functions \tilde{m}_x and $\tilde{\sigma}_x^2$ correspond with the shape of the histogram in Fig. 6.

Fig .7: Approximated membership functions \widetilde{m}_x a \widetilde{s}_x^2 3.2 $\mu(m_{\star})$ \widetilde{m}_{r} $\alpha - \check{r}e$ $\mu(s_x^2)^2$ α -hladina α -hladina 1,00 0,00 1,00 0,00 2,07 3,06 0,75 0,93 1,56 0,50 0,75 0,50 3,22 0,79 1,92 0,50 0,75 0,75 0,50 3.38 0,25 2.4 .42 0,90 0,67 0,25 0,90 2.08 3,90 0.99 0,51 0,00 0,99 0,00 2,0 3,0 2,89 3,5 0,0 ż,0 m 3,0 Source: own creation

3.2.1 3. 2 Unemployment Real Data Set Analysis

In order to accomplish the fuzzy stochastic analysis of the real random variable, it is necessary to gain a selection statistical set which would simultaneously be considered vague. Such a set would be the observations shown in Tab. IV. which represent the unemployment rate values in the USA in the years 1985 to 2008 (Statistical Information, 2013). This data cannot be statistically correct, as it was mined over a long period in which it was impossible to create duplicable conditions and, at the same time, it is small (n = 24).

Tab. IV: Real random file

| [7.0 6.6 5.7 5.3 5.4 6.3 7.3 7.4 6.5 5.5 5.6 5.4 4.7 4.4 | |
|--|--|
| 4.0 3.9 5.7 6.0 5.7 5.4 4.9 4.4 5.0 7.4] | |
| Source: own creation | |

The statistical parameters are the median med = 5.5500, and the mean value $\overline{m}_x = 5.6458$. Fig 8. shows the histogram of this set.

Fig.8: Histogram of real random file



Source: own creation

Exploratory graphic normality test method (using the normal probability graph method) was used to estimate the distribution type of the set (Budíková, Králová and Maroš, 2010). The result (Fig. 9) confirms the presumption of normality.





Fig. 10a shows the result of the determination of the membership function of the median fuzzy number using a histogram. This approach would have to be used in the scenario of an unknown distribution of the fuzzy random variable. In Fig. 10b, the approximated triangular membership function is normed.



Fig. 10a: Membership function approximation



Source own creation

The results of the determination of the membership function of the fuzzy number of the median and sparsity via their confidence intervals are shown in the Tab. V. and Tab. VI. The corresponding functions are shown in Fig. 11.

| Tab. V: Membership function | <i>m</i> , |
|-----------------------------|------------|
|-----------------------------|------------|

| Confidence level a | Membership function α-cut level | L- boundary | R-boundary |
|-----------------------|------------------------------------|-------------|------------|
| 0,99 | 0,00 | 5,06 | 6,23 |
| 0,90 | 0,25 | 5,29 | 6,00 |
| 0,75 | 0,50 | 5,40 | 5,90 |
| 0,50 | 0,75 | 5,50 | 5,79 |
| | 1,00 | 5, | 64 |

Source: own creation

Tab.VI: Membership function \tilde{s}_x^2

| Confidence level a | Membership function α-cut level | L- boundary | R-boundary |
|-----------------------|------------------------------------|-------------|------------|
| 0,99 | 0,00 | 0,54 | 2,59 |
| 0,90 | 0,25 | 0,68 | 1,83 |
| 0,75 | 0,50 | 0,77 | 1,54 |
| 0,50 | 0,75 | 0,88 | 1,32 |
| | 1,00 | 1,1 | 04 |

Source: own creation



Fig.11: Approximated membership functions \widetilde{m}_{r} and \widetilde{s}_{r}^{2}

Source: own creation

The mean value \tilde{m}_{x} and sparsity \tilde{s}_{x}^{2} of the fuzzy random value are therefore defined by the shape of their membership function, where this function bears the information about the size and uncertainty of the values. If such fuzzy mean value or fuzzy sparsity is to be used in further mathematical calculations, the solution will be determined not using the laws of normal, but rather of the fuzzy branch of mathematics (Morderson and Nair, 2001). The result of such calculation is not a regular, but rather a fuzzy number, once again bearing the information about its value as well as its uncertainty. The uncertainty, the rate of witch is the width of the fuzzy interval, is not a stochastic uncertainty caused by the effects of random influence (mean value sparsity). It is caused by vagueness, witch is in turn caused by our lack of information about its size (the verbal equivalent of "probably, approximately, roughly").

The example is the determination of the amount of people out of the unemployed, who will in the given year find work. Let us suppose that on January 1st, there will be 5 million employed and 0.5 million unemployed individuals. It is also supposed that the year will see 38 % of the unemployed finding a job and the real unemployment rate will be equal to the natural rate (Soukup et. al., 2009). Let us consider that the fuzzy number \widetilde{Z} of the amount of people in the productive age has the membership function according to Fig. 12. The amount of unemployed in the year can be calculated using the fuzzy arithmetic (Keprt, 2012) and the following relation

$$\widetilde{N} = \widetilde{m}_{r}.\widetilde{Z}$$

where \overline{m}_{x} = 5,64% is the fuzzy natural unemployment rate, \overline{Z} = 5,5 mil people is the fuzzy amount of people in the productive age (Tab. VII.).

| Confidence level a | Membership function α-cut level | L- boundary | R-boundary |
|--------------------|------------------------------------|-------------|------------|
| 0,99 | 0,00 | 26,80 | 35,50 |
| 0,90 | 0,25 | 28,30 | 33,90 |
| 0,75 | 0,50 | 29,16 | 33,04 |
| 0,50 | 0,75 | 29,97 | 32,13 |
| | 1,00 | 31 | .02 |

Tab.VII: Membership function \widetilde{N}

Source: own creation

The amount of people able to find a work position can be calculated using the fuzzy arithmetic in the following manner

 $\widetilde{P} = n \cdot \widetilde{N}$

Where *n* = 38 % is the rate of finding the employment, \tilde{N} is the fuzzy amount of the unemployed (Tab. VIII.).

| Confidence level α | Membership function α-cut level | L- boundary | R-boundary |
|-----------------------|---------------------------------------|-------------|------------|
| 0,99 | 0,00 | 10,18 | 13,50 |
| 0,90 | 0,25 | 10,75 | 12,88 |
| 0,75 | 0,50 | 11,08 | 12,55 |
| 0,50 | 0,75 | 11,39 | 12,21 |
| | 1,00 | | 11,79 |

Tab.VIII: Membership function \widetilde{P}

Source: own creation

The shapes of the membership functions of the fuzzy numbers of the calculations results are stated in Fig. 13 and Fig. 14.





The cores of the triangular fuzzy sets state the clear value of the result, whilst the fuzzy interval states the rate of uncertainty of the result. The level of possibility of occurrence of a value from an uncertain interval is given by the level of membership of this value to a fuzzy set.

Fig.13: Approximated membership





Source: own creation

The uncertainty of the results, given by the width of the fuzzy intervals, provides the user with the information about the vagueness of the resulting values. This information is in the given concept interesting and can be, by an expert, well interpreted.

4 Conclusion

Statistical methods can only reflect the uncertainty of the stochastic type. Inaccurate, unreliable data, uncertainties which cannot be described or are insufficiently statistically described can be taken into account merely in an approximate fashion. Thus, conventional methods of the statistical analysis can be used only in a limited scope.

Application of the fuzzy stochastic approaches for the analysis of selection files can be tangible in two cases. It is tangible especially when the test does not prove the validity of the hypothesis about the stochastic character of the selection file. Secondly, it can be tangible in the scenario when the conditions of representativeness of the selection file are not well met. This means that the file is of a small scale, the statistical data are of uncertain accuracy or were gained in uncertain or non-reproduceable conditions.

A number of analyses of practical systems in the area of social sciences require the use of an integrated – fuzzy stochastic – approach and thus formalize the random fuzzy variables. However the fuzzy random values may partially reflect a stochastic character, they may not be without any doubt processed by purely statistical methods, as their stochasticity is accompanied and interrupted by fuzzitivity. A fuzzy random value can be understood as a random value which was measured under uncertain conditions, i.e. the observation was not done under exactly defined experimental conditions. The processing of fuzzy stochastic uncertainty makes use of the approaches of the theory of fuzzy random values. It uses mainly objective information, but subjective information can also be used.

The paper contains the issue of the probability theory of fuzzy random variables and determination of their numerical fuzzy characteristics. The paper also introduces the definition of the most important terms and gives examples of the use of fuzzy stochastic analysis of files of economic data.

Acknowledgements

This paper was created with the financial support and in the framework of the Czech Science Foundation project no. P403/12/1811: Unconventional Managerial Decision-Making Methods Development in Enterprise Economics and Public Economy.

5 Bibliography

- [1] BUDÍKOVÁ, M., KRÁLOVÁ, M. and MAROŠ, B. 2010. *Průvodce základními statistickými metodami*. Praha: Grada Publishing. ISBN 978-80-247-3243-5.
- [2] KEPRT, A. 2012. Programový systém pro fuzzy aritmetiku s využitím alfa-řezů. In *Ekonomika-Management-Inovace*. Olomouc: MVŠO. 4(3). ISSN 1804-1299.
- [3] MORDESON, J. N. and NAIR, P. S. 2001. *Fuzzy Mathematics*, Physica-Verlag. ISBN 3-7908-1420-2.
- [4] MÖLLER, B. and BEER, M. 2004. Fuzzy Randomness Uncertainty in Civil Engineering and Computational Mechanics. Springer. ISBN 3-540-40294-2.
- [5] MÖLLER, B. 2004. Fuzzy Randomness A Contribution to Imprecise Probability. ZAMM Z Agnew. Match. Mech.84. No.10-14: 754-764.WILLEY-VCH.
- [6] NOVÁK, V., PERFILIEVA, I. and DVOŘÁK, A. 2016. Insight Into Fuzzy Modelling. Wiley. ISBN 978-1-119-19318-0.
- [7] SIEGEL, A. F. 2012. Practical Bussiness Statistics. Elsevier Inc. ISBN 978-0-12-385208-3.
- [8] SOUKUP, J. et. al. 2009. Makroekonomie: cvičebnice. Melandrium. ISBN 978-80-86175-64-5.
- [9] STATISTICAL INFORMATION. *Czech Statistical Office*. [online]. [cit. 2013-05-12]. Dostupné z: WWW: http://www.czso.cz/csu/redakce.nsf/i/cr:_makroekonomicke_udaje/\$File/HLMAKRO.xls
- [10] STATISTICA Version 12. 2014. StatSoft Releases Version 12 of STATISTICA Software. [online]. [cit. 2014-10-06]. Dostupné z: http://www.statsoft.com/v12