



Claim reserving with fuzzy regression

Tahereh BAHRAMI^{1,*}, Masuod BAHRAMI²

¹*Behbahan Khaatam Alanbiyaa University of Technology*

²*Yasouj Branch, Islamic Azad University, Yasouj, Iran,*

Received: 01.02.2015; Accepted: 05.05.2015

Abstract. Claims reserving plays a key role for the insurance. Therefore, various statistical methods are used to provide for an adequate amount of claim reserves. Since claim reserves are always variable, fuzzy set theory is used to handle this variability. In this paper, non-symmetric fuzzy regression is integrated in the Taylor's method to develop a new method for claim reserving.

Keywords: Fuzzy number, fuzzy regression, claim reserves, Taylor's geometric separation method

1. INTRODUCTION

Predicting claim provisions is an important issue in statistics. There exist several different classifications among the methods of claim reserving. The highest level of classification is the division of methods into two main categories: classic and stochastic methods. The classic method uses a deterministic approach and only yields the fair value of future claims. The stochastic method yields better predictions in comparison to the classic method. Furthermore, variability of the claims is also included in it.

In this paper we use the second method and In addition to pricing the claim reserves, probable variations are dealt with. Non-symmetric fuzzy regression is integrated in the Taylor's geometric separation method. Fuzzy regression is derived from fuzzy sets theory.

Fuzzy sets theory is used for problems with inadequate and ambiguous data. That is the reason Ramnasy, Doxtin and Fedrisy take advantage of fuzzy regression for economical analysis.

2. FUZZY NUMBER

2.1. Grade of membership and membership function

A fuzzy subset A , defined in the universal space X , is a function defined in X which assumes values in the range $[0, 1]$.

$$\tilde{A}: X \rightarrow [0, 1] \quad (1)$$

Where $x \in X$ and $\mu_{\tilde{A}}(x)$ is the value of the function for this element. This value is called the grade of membership.

*Corresponding author. Email address: Bahrami@bkatu.ac.ir

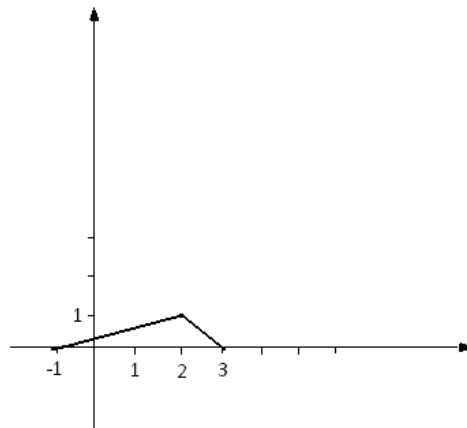
2.2. Triangular fuzzy number

Fuzzy numbers are used in the present paper, since the calculation and interpretation of triangular fuzzy numbers is quite simple. A triangular fuzzy number is denoted by $\tilde{A} = (a, l_a, r_a)$, which a, l_a, r_a are the center value, right spread and left spread respectively. For instance, an actuary's expectation of the inflation rate of the cost of the claims to be about 2 percent with a variance of less than 1 percent, is denoted as $(0.02, 0.01, 0.01)$. triangular fuzzy number with membership function $\mu_{\tilde{A}}(x)$ is denoted as follows:

$$\mu_{\tilde{A}}(x) = \begin{cases} \frac{x - a + l_a}{l_a} & \text{if } a - l_a < x \leq a \\ \frac{x - a + r_a}{r_a} & \text{if } a < x \leq a + r_a \\ 0 & \text{otherwise} \end{cases} \quad (2)$$

2.3. Example

The triangular fuzzy number for $a = 2, l_a = 1, r_a = 3$ is illustrated as follows:



Graph1. For example 2.3

Note: Whenever the right and left spreads of a triangular fuzzy number are equal, then the triangular fuzzy number is called symmetric and denoted by (a, l_a) , otherwise it is a non-symmetric number.

3. FUZZY REGRESSION WITH NON-SYMMETRIC COEFFICIENTS

As with ordinary regression, the goal of fuzzy regression is determining the relation between a dependent variable and a set of independent variables. Suppose that the dependent variable is a linear combination of independent variables, then this relation is obtained from a sample of n observations $\{(Y_1, X_1), (Y_2, X_2), \dots, (Y_i, X_i), \dots, (Y_n, X_n)\}$ where X_j is the j 'th observation of the dependent variable and X_j is m -dimensional: $X_j = (X_{1j}, X_{2j}, \dots, X_{ij}, \dots, X_{mj})$.

Furthermore, X_{ij} is the observed value of the i 'th variable in the j 'th observation Y_j is the j 'th observation of the above mentioned variable which is always a crispy number.

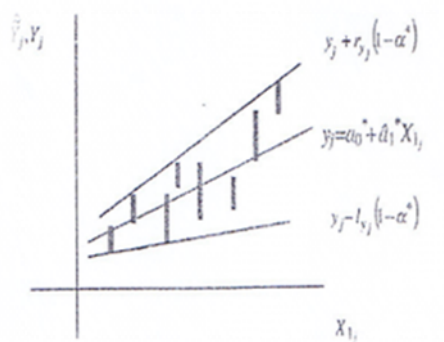
Now the linear fuzzy function is estimated as follows:

$$\hat{\tilde{Y}}_j = \tilde{A}_0 + \tilde{A}_1 X_{1j} + \dots + \tilde{A}_m X_{mj} \tag{3}$$

Where $\hat{\tilde{Y}}_j$ is the approximation of the Y_j using fuzzy number after setting $\tilde{A}_0, \tilde{A}_1, \tilde{A}_2, \dots, \tilde{A}_m$.

3.1. Example

Fuzzy linear regression approximation for $\hat{\tilde{Y}}_j = \tilde{A}_0 + \tilde{A}_1 X_{1j}$, $j = 1, 2, 3, \dots, n$ is illustrated below. Bold black lines represent dependent variables.



Graph 2. for example 3.1

4. RUN-OFF TRIANGLE

In non-insurance applications, run-off triangle is used for predicting the number of future claims and the delay in their payments, as well as for non-criminal.

Examples of application of the run-off triangle:

- a) Annual reports (such as accumulated claims reported in year x)
- b) Annual bills (such as traffic tickets issued in a year)

In order to estimate the future claims some amount of money should be reserved for annual accidents. Anyhow, the basis of commitment in run-off triangle for estimating the solution for the reserved claims is the immediate payment for each accident.

Table 1. Run-off triangle

		development period						
		K	k-1	j	
Origin period	0	$C_{0,0}$	$C_{0,1}$:	$C_{0,j}$:	$C_{0,k-1}$	$C_{0,k}$
	1	$C_{1,0}$	$C_{1,1}$:	$C_{1,j}$:	$C_{1,k-1}$	
	:	:	:	:	:	:		
	i	$C_{i,0}$	$C_{i,1}$:	$C_{i,j}$			
	:	:	:	:				
	k-1	$C_{k-1,0}$	$C_{k-1,1}$					

	k	$C_{k,0}$					
--	---	-----------	--	--	--	--	--

Where $C_{i,j}$, $j = 1,2,3,\dots,n$ is the cost of the claims for all insurances issued in the i 'th origin period and the j 'th claim period.

It is evident from table 1 that:

In the first origin period, the first row of the triangle, none of the claims are settled, i.e this is the development period 0.

In the second origin period, claims of one period are settled and there are $k-1$ periods to be settled, so this is the first development period.

And finally, in the k 'th origin period, all claims are settled and this is the k 'th development period, i.e the vertex of the triangle represents the peak of company's development.

4.1. About Taylor's geometric separation method

We propose an extension of Taylor's geometric separation method for fuzzy regression (1997). In this method it is assumed that the number of observations has fluctuation from one origin period to another. Therefore, the aim of this method is to estimate the incurred but not reported reserves. So the insurer should always take account of this type of reserves. Estimation of the reserves is usually carried out based on the past observations or is planned considering the monthly incidents and circumstances. Therefore IBNR (incurred but not reported reserves) varies from month to month. Considering the data in table 1, the average cost for each of the claims in the i 'th origin period in j 'th development period is obtained as follows

$$\tilde{S}_{i,j} = \frac{C_{i,j}}{\tilde{N}_i} \tag{5}$$

Where N_i is the number of the claims in the i 'th origin period.

In Taylor's method it is assumed that $S_{i,j}$, $i, j = 1,2,3,\dots,k$ could be divided into two factors: first factor, P_i , the relative significance of the i 'th development factor in comparison to the origin number and the second factor, π_{i+j} , the impact of increase in the cost of claims in the $i+j$ 'th calendar period. It is inferred from Taylor's method that $S_{i,j}$, $i, j = 1,2,3,\dots,k$ is in fact the multiplication of both factors:

$$S_{i,j} = P_i \pi_{i+j} \quad i, j = 1,2,3,\dots,k \tag{6}$$

Therefore, parameters P_i and π_{i+j} , $i + j = k$ are simply obtained using linear regression.

In Taylor's geometric separation method it is assumed that the number of observations in the i 'th origin period is the fuzzy number $\tilde{N}_i = (n_i, l_{n_i}, r_{n_i})$, $i = 1,2,\dots,n$, where n_i , l_{n_i} and r_{n_i} are the number of observations for occurrence period, number of outstanding claims and number of claims in the development period, respectively. Consequently, the average cost which is now a triangular fuzzy number is obtained by the equation below:

$$\tilde{S}_{i+j} = \frac{C_{i,j}}{\tilde{N}_i} \tag{7}$$

Therefore:

$$\hat{S}_{i,j} = \tilde{P}_j \tilde{\pi}_{i,j} \tag{8}$$

Which can be denoted as:

$$\ln \hat{S}_{i,j} = \ln \tilde{P}_j + \ln \tilde{\pi}_{i,j}$$

(9)

$$\ln \hat{S}_{i,j} = (\ln P_j, l_{\ln P_j}, r_{\ln P_j}) + (\ln \pi_{i+j}, l_{\ln \pi_{i+j}}, r_{\ln \pi_{i+j}})$$

it holds that:

(10)

Now assuming

$$\ln \tilde{P}_j = (\ln \tilde{P}_j, l_{\ln \tilde{P}_j}, r_{\ln \tilde{P}_j}) \text{ and}$$

$$\ln \tilde{\pi}_{i+j} = (\ln \pi_{i+j}, l_{\ln \pi_{i+j}}, r_{\ln \pi_{i+j}})$$

Which means the average cost is a triangular fuzzy number and easy to deal with. To adjust the equation (10) the center values for $\ln \tilde{P}_j$ and $\ln \tilde{\pi}_{i+j}$ should be calculated using the non-symmetric regression explained in the previous section.

5- Estimating the cost of the claims

$\tilde{S}_{i,j}$ Is not a triangular fuzzy number but could be approximated to the triangular fuzzy number $\hat{S}_{i,j} \approx (\hat{S}_{i,j}, l_{\hat{S}_{i,j}}, r_{\hat{S}_{i,j}})$ where:

$$\hat{S}_{i,j} = e^{\ln \hat{S}_{i,j}} \tag{11}$$

Since $\hat{C}_{i,j} = \hat{S}_{i,j} \tilde{N}_i$ is a triangular fuzzy number therefore $PRO_i = \sum_{j=n-i+1}^n \hat{C}_{i,j}$ is also a triangular fuzzy number:

$$PRO^* = (pro, l_{pro}, r_{pro}) = \left(\sum_{i=1}^n pro_i, \sum_{i=1}^n l_{pro_i}, \sum_{i=1}^n r_{pro_i} \right) \tag{12}$$

$$= \left(\sum_{i=1}^n \sum_{j=n-i+1}^n \hat{c}_{i,j}, \sum_{i=1}^n \sum_{j=n-i+1}^n l_{\hat{c}_{i,j}}, \sum_{i=1}^n \sum_{j=n-i+1}^n r_{\hat{c}_{i,j}} \right)$$

Now to convert the cost of the claims into a crispy number, a number called the expected value of Compos Gonzalles with risk averse β , ($0 \leq \beta \leq 1$) is used, which is denoted for the fuzzy number \tilde{A} as follows:

$$EV[\hat{A}, \beta] = (1 - \beta) \int \underline{A}(\alpha) d\alpha + \beta \int \bar{A}(\alpha) d\alpha \tag{13}$$

If $\tilde{A} = (a, l_a, r_a)$ then:

$$EV[\tilde{A}, \beta] = a - \frac{1 - \beta}{2} l_a + \frac{\beta}{2} r_a \tag{14}$$

Then PRO should be converted to the thread number PRO^* using (8) and (9):

$$\begin{aligned}
 PRO^* &= EV[P\tilde{R}O, \beta] = pro - \frac{1-\beta}{2} l_{pro} + \frac{\beta}{2} r_{pro} = \\
 &\sum_{i=1}^n \sum_{j=n-i+1}^n \hat{c}_{i,j} - \frac{(1-\beta)}{2} \sum_{i=1}^n \sum_{j=n-i+1}^n l_{\hat{c}_{i,j}} + \frac{\beta}{2} \sum_{i=1}^n \sum_{j=n-i+1}^n r_{\hat{c}_{i,j}}
 \end{aligned} \tag{15}$$

5. CONCLUSION

There is no need for lots of data when adjusting claim reserves of the insurance. In this regard fuzzy sets theory could be useful. In the present paper we have integrated the fuzzy sets theory into the Taylor's method, not only as a way for estimating incurred but not reported reserves but also for outstanding reserves. The method proposed in the present paper has two advantages. First, the estimates obtained after developing the fuzzy regression models is a triangular fuzzy number which is easy to deal with. Second, estimate regarding the variability of the reserves is also a fuzzy number.

REFERENCES

- [1] Chang, P.T., Lee, E.S., 1996. A generalization fuzzy weighted least-squares regression. *Fuzzy Sets systems* 82, 289-298.
- [2] Chang, P.T., 1997. Fuzzy seasonality forecasting. *Fuzzy Sets Systems* 90, 1-10.
- [3] Diamond, P., Kloeden, P. 1994. *Metric Spaces of Fuzzy Sets. Theory and Application.* World Scientific, Singapore.
- [4] Dubois, D., Prade, H., 1980. *Fuzzy Sets and Systems: Theory and Applications.* Academic Press, New York.
- [5] Kacprzyk, J., Fedrizzi, M, (Eds.), 1992. *Fuzzy regression Analysis.* Omnitech and Heidelberg: Physic, Waesaw.
- [6] Tanaka, H., S. Uejima and K. Asai. 1982. Liner regression analysis with Fuzzy model. *IEEE Trans. System Man Cybernet.* 12: 903-907.
- [7] Zimmermann, H. J. 1996. *Fuzzy Sets Theory and Its application.* Kluwer Academic Pub. , Boston.