Possibilistic Modeling for Loss Distribution and Premium Calculation

Zhen Huang
Otterbein College
Westerville, OH 43081

Lijia Guo
Algorithmics, Fitch Group
Newton, MA 02466

ABSTRICT

This paper uses the possibility distribution approach to estimate the insurance loss amount. A special class of parametric possibility distributions is used to model insurance loss variables. The parameters of the possibility distribution are estimated by combining statistical analysis of sample data and domain knowledge provided by actuarial experts. Insurance premiums are calculated using possibilistic mean and possibilistic variation. Estimation of possibility distribution of aggregate loss amount is also discussed.

Key word: insurance loss distribution, premium, possibility distribution, possibilistic mean, possibilistic variation, and aggregate loss amount

1. Introduction

Conventionally the uncertainty of potential insurance loss amount is modeled by probability theory. The insurance premium is then calculated by using the estimated probability distribution of probabilistic loss variable. Possibility theory provides alternative way for modeling uncertainty [Zimmermann (2001)]. Possibility theory, originated from fuzzy sets theory [Zadeh (1965)], was first proposed by Zadeh (1978) and was further developed by Dubois and Prade (1988). The idea of using possibility distribution to model uncertainty has been applied to different problems. One of such applications is to model the rates of return in portfolio selection problems by possibility distributions, rather than conventional probability distributions [Guo and Huang (1996), Inuiguchi and Tanino (2000), Verchera, Bermúdeza and Segurab (2007)]. Some of the applications of possibility theory and fuzzy logic in insurance and financial risk management were reviewed by Shapiro (2004). In this paper, we suggest using possibility distribution to model insurance loss amount. Our paper is organized as follows. In Section 2 we briefly introduce some concepts on possibility distributions. In Section 3 we give a hybrid procedure of estimating possibility distribution of loss variable by using sample data and experts' knowledge. In Section 4 we provide formulas for measuring possibilistic mean and possibilistic variation. Some premium calculation methods for possibilistic loss variables are proposed. Discussion of possibility distribution, possibilistic mean, and possibilistic variation of aggregate loss claim amount is given in Section 5. Conclusions are given in Section 6.

2. Possibility Distribution

Definition 2.1 [Dubois and Prade (1980), Zimmermann (2001)]
Possibility distribution π_X of a variable X on real number set $\mathbb{R}$ is a mapping from $\mathbb{R}$ to interval [0, 1], and there is at least one $c_0 \in \mathbb{R}$ so that $\pi_X(c_0)=1$.

Possibility distribution can be represented by fuzzy number using membership function as possibility function.

Example 1:
A triangular possibility distribution $\pi_X = (m, s_l, s_r)_{LR}$ has possibility function

$$\pi_X(x) = \begin{cases} 1-(m-x)/s_l & \text{if } x \in [m-s_l, m), \\ 1 & \text{if } x = m, \\ 1-(x-m)/s_r & \text{if } x \in (m, m+s_r]. \end{cases}$$

Definition 2.2
A LR-possibility distribution $\pi_X = (m_l, m_r, s_l, s_r)_{LR}$ has possibility function

$$\pi_X(x) = \begin{cases} L\left(\dfrac{m_l - x}{s_l}\right) & \text{if } x \in [m_l - s_l, m_l), \\ 1 & \text{if } x \in [m_l, m_r], \\ R\left(\dfrac{x - m_r}{s_r}\right) & \text{if } x \in (m_r, m_r + s_r], \end{cases}$$

where the left reference functions $L:[0,1] \to [0,+\infty)$ and the right reference function $R:[0,1] \to [0,+\infty)$ satisfy the following conditions:
(1) $L(0) = R(0) = 1$,
(2) $L(x)$ and $R(x)$ are strictly decreasing,
(3) $L(x)$ and $R(x)$ are upper semi-continuous on bounded $\text{supp}(\tilde{A}) = \{x \in \mathbb{R} : \pi_X(x) > 0\}$.
The interval $[m_l, m_r]$ is peak of π_X, m_l and m_r are left-mode and right-mode, and s_l and s_r are left-spread and right-spread.

Triangular possibility distribution in Example 1 is a *LR*-possibility distribution with $m_l = m_r = m$, and the left and right reference functions $L(u) = R(u) = 1 - u,\ 0 \le u \le 1$.

A class of specially constructed *LR*-possibility distribution that can be used for modeling insurance loss variables is given in Huang and Guo (2007).

3. Possibilistic Modeling of Loss Variable

In the following, we use possibility distribution π_X to model insurance loss variable *X* rather than the conventional probability distribution approach. For every $x \in \mathbb{R}$, $\pi_X(x)$ is the possibility that *x* can possibly be assigned to loss variable *X*. The possibility distribution and probability distribution for same variable *X* must satisfy the Probability/Possibility Consistency Principle proposed by Zadeh (1978):

The possibility of an event is always greater than or equal to its probability.

Mathematically, let π_X be the possibility distribution and p_X be the probability density function for the same variable *X*, then the possibility/probability consistency principle is

$$\sup_{x \in D} \pi_X(x) = \text{Possibility}(D) \ge \text{Probability}(D) = \int_D p_X(x)dx\,,$$

for any union D of disjoint intervals [Dubois and Prade (1980)].

In insurance, loss variables typically have light or heavy right tails. We propose a class of parametric functions, called power-exponential mixed functions, to fit histogram of loss amount sample data.

Definition 3.1
The class of power-exponential mixed functions is defined as

$$\hbar(x) = \begin{cases} a\left[1 - \left(\dfrac{m_l - x}{s_l}\right)^{p_l}\right] & \text{if } x \in [m_l - s_l, m_l), \\ a & \text{if } x \in [m_l, m_r], \\ a\left[\exp\left(-p_r\left(\dfrac{x - m_r}{s_r}\right)\right)\right] & \text{if } x \in (m_r, m_r + s_r], \\ 0 & \text{otherwise.} \end{cases}$$

where $a > 0$, $p_l \ge 1$, and $p_r > 0$

Note that $\hbar(x)$ has seven parameters: height parameter a; range parameters m_l, m_r, s_l, and s_r; and sharp parameters p_l and p_r. If we normalize $\hbar(x)$ through dividing $\hbar(x)$ by $\sup \hbar$, then we obtain a power-exponential mixed *LR*-possibility distribution

$$\pi_X = (m_l, m_r, s_l, s_r)_{LR} = \frac{\hbar(x)}{\sup \hbar} = \frac{\hbar(x)}{a}, \text{ for } x \in \mathbb{R}.$$

Since $\hbar(x)$ has seven parameters, by choosing different parameter values $\hbar(x)$ can be used to fit a large class of loss variables either with light or heavy right tails. More importantly, we have proved that the possibility distribution and the probability distribution induce form $\hbar(x)$ satisfy the Possibility/Probability Consistency Principle [Huang and Guo (2007)]:

The possibility distribution $\pi_X(x) = \dfrac{\hbar(x)}{\sup \hbar} = \dfrac{\hbar(x)}{a}$, for $x \in \mathbb{R}$, and the probability density function $p_X(x) = \hbar(x) / \int_{-\infty}^{\infty} \hbar(x)dx$ satisfy the *Possibility-Probability Consistency Principle*, namely for any union D of disjoint intervals

$$\sup_{x \in D} \pi_X(x) = \text{Possibility}(D) \geq \text{Probability}(D) = \int_D p_X(x)dx.$$

Our procedure uses sample data combining with actuarial experts' knowledge to estimate seven parameters of $\hbar(x)$. Then the possibility distribution $\pi_X(x)$ for loss variable X is induced from $\hbar(x)$. A detailed description of the hybrid procedure and numerical examples are given by Huang and Guo (2007).

4. Premium Calculation for Possibilistic Loss Variable

Dubois and Prade (1987) first introduced the mean value of a fuzzy number as a closed interval bounded by the expectations calculated from its upper and lower distribution functions. Carlsson and Full'er (2001) defined crisp possibilistic mean value and variance of fuzzy numbers. Majlender and Full'er (2003) introduced weighted possibilistic mean and variance of fuzzy numbers. Since possibility distributions can be represented as fuzzy numbers, the concepts of possibilistic mean value and variance for fuzzy numbers can be naturally translated to possibility distributions. In this section all possibility distributions used are assumed to be *LR*-possibility distributions. We need to introduce the concept of α-level cut for possibility distribution π_X before defining possibilistic mean.

Definition 4.1
α-level cut of a possibility distribution π_X is define by

$$(\pi_X)_\alpha = \{u \in \mathbb{R} \mid \pi_X(u) \ge \alpha\}, \qquad \alpha > 0.$$

It is easy to see that for *LR*-possibility distribution, $(\pi_X)_\alpha = [\inf(\pi_X)_\alpha, \sup(\pi_X)_\alpha]$.

Definition 4.2 (Carlsson and Full'er, 2001)

(a) The possibilistic mean of π_X is defined by

$$\bar{M}(\pi_X) = \int_0^1 \alpha\left(\sup(\pi_X)_\alpha + \inf(\pi_X)_\alpha\right) d\alpha.$$

(b) The possibilistic variance of π_X is defined by

$$Var(\pi_X) = \frac{1}{2}\int_0^1 \alpha\left(\sup(\pi_X)_\alpha - \inf(\pi_X)_\alpha\right)^2 d\alpha.$$

(c) The possibilistic standard deviation of π_X is defined by

$$\sigma_{\pi_X} = \sqrt{Var(\pi_X)} = \sqrt{\frac{1}{2}\int_0^1 \alpha\left(\sup(\pi_X)_\alpha - \inf(\pi_X)_\alpha\right)^2 d\alpha}\ .$$

Possibilistic standard deviation σ_{π_X} defined above can be interpreted as a measure of variation for the possibility distribution π_X. Since $\left(\sup(\pi_X)_\alpha - \inf(\pi_X)_\alpha\right)$ in definition of σ_{π_X} is always greater or equal to zero for any $\alpha \in [0,1]$, we propose another way to measure the variation of possibility distribution π_X. Our following definition of possibilistic variation is similar to the absolute deviation in probability theory.

Definition 4.3

Possibilistic variation of π_X is defined by

$$\bar{V}(\pi_X) = \int_0^1 \alpha(\sup(\pi_X)_\alpha - \inf(\pi_X)_\alpha) d\alpha.$$

Since $(\pi_X)_\alpha = [\inf(\pi_X)_\alpha, \sup(\pi_X)_\alpha]$, the difference $(\sup(\pi_X)_\alpha - \inf(\pi_X)_\alpha)$ can be viewed as variation of $(\pi_X)_\alpha$, the α-level cut of π_X. It follows that the possibilistic variation $\bar{V}(\pi_X)$ is the expected value of variations of its level cuts. $\bar{V}(\pi_X)$ can be used as a measure for variation of π_X.

Example 4.1

Let $\pi_X = (m, s_l, s_r)_{LR}$ be a triangular possibility distribution, then

(a) $\quad (\pi_X)_\alpha = \left[m - s_l(1-\alpha), m + s_r(1-\alpha)\right], \qquad$ for $\alpha \in [0,1]$.

(b) The possibilistic mean value of π_X is

$$\bar{M}(\pi_X) = m + \frac{s_r - s_l}{6}.$$

(c) The possibilistic variation of π_X is

$$\bar{V}(\pi_X) = \frac{s_l + s_r}{6}.$$

Remarks:

(1) If $s_r = s_l$ (possibility function is symmetric) then $\bar{M}(\pi_X) = m$.

(2) The range of $\pi_X = (m, s_l, s_r)_{LR}$ is $s_l + s_r$, it means that $\bar{V}(\pi_X) = \frac{s_l + s_r}{6} = \frac{range}{6}$. This is similar to the estimation formula for standard deviation of random variable *X* in statistics $\sigma_X \approx \frac{range}{6}$.

Our premium calculation for possibilistic loss variable is analogous to the conventional variance or standard deviation loading premium calculation method when the loss variable is modeled by probability distribution [Cizek, Hardle, and Weron (2005)].

We proposed the following risk loading premium calculation method.

$$P_{RL}(\theta) = \bar{M}(\pi_X) + \theta\bar{V}(\pi_X),\ \theta \geq 0,$$ where θ is the risk loading factor.

A detailed discussion of other premium calculation methods for possibilistic loss variables are given by Huang and Guo (2007).

5. Possibilistic Aggregate Loss Variable

Addition and scalar multiplication of possibilistic variables distribution can be defined by the sup-min extension principle as addition and scalar multiplication of *LR*-fuzzy numbers [Dubois and Prade (1980)].

Addition of possibilistic variables with LR-possibility distributions

Let $X_i, i = 1, 2,$ be two possibilistic variables with *LR*-possibility distributions $\pi_{X_i} = (m_{l_i}, m_{r_i}, s_{l_i}, s_{r_i})_{LR}$, $i = 1, 2.$ Then the possibility distribution of $X_1 + X_2$ is a *LR*-possibility distribution $\pi_{X_1+X_2} = (m_{l_1} + m_{l_2}, m_{r_1} + m_{r_2}, s_{l_1} + s_{l_2}, s_{r_1} + s_{r_2})_{LR}$.

Scalar multiplication

Let *X* be a possibilistic variable with *LR*-possibility distribution $\pi_X = (m_l, m_r, t_l, t_r)_{LR}$, and *w* a nonnegative real number. Then the possibility distribution of scalar multiplication, *wX*, is a *LR*-possibility distribution $\pi_{wX} = (wm_l, wm_r, wt_l, wt_r)_{LR}$.

Proposition 5.1

Let $X_1, X_2, ..., X_n$ be n possibilistic variables and $w_1, w_2, ..., w_n$ be n nonnegative real numbers. Each X_i has *LR*-possibility distribution $\pi_{X_i} = (m_{li}, m_{ri}, s_{li}, s_{ri})_{LR}$ with the left reference function L_{X_i} and the right reference function R_{X_i}, $i = 1, 2, ..., n$. Then the possibility distribution of the linear combination $X = \sum_{i=1}^{n} w_i X_i$ is also a *LR*-possibility distribution $\pi_X = \left(\sum_{i=1}^{n} w_i m_{li}, \sum_{i=1}^{n} w_i m_{ri}, \sum_{i=1}^{n} w_i s_{li}, \sum_{i=1}^{n} w_i s_{ri} \right)_{LR}$ with some left and right reference functions L_X and R_X.

Proposition 5.2

Let $X = \sum_{i=1}^{n} w_i X_i$, then

(1) $\bar{M}(\pi_X) = \sum_{i=1}^{n} w_i \bar{M}(\pi_{X_i})$.

(2) $\bar{V}(\pi_X) = \sum_{i=1}^{n} w_i \bar{V}\left(\pi_{X_i}\right)$.

Furthermore if all π_{X_i}, $i = 1, 2, ..., n$, have the same the left reference function L and the same right reference function R, then the left and right reference functions of π_X are the same as L and R of π_{X_i}. Combining this with Proposition 5.1, we give a complete description of possibility distribution of π_X in the following.

Proposition 5.3

Let $X_i, i = 1, 2, ..., n,$ be n possibilistic variables with *LR*-possibility distribution $\pi_{X_i} = (m_{li}, m_{ri}, s_{li}, s_{ri})_{LR}$ and w_i, $i = 1, 2, ..., n,$ be n nonnegative real numbers. If all $\pi_{X_i}, i = 1, 2, ..., n,$ have the same left reference function L and the same right reference function *R*, then the possibility distribution of the linear combination $X = \sum_{i=1}^{n} w_i X_i$ is also a *LR*-possibility distribution $\pi_X = \left(\sum_{i=1}^{n} w_i m_{li}, \sum_{i=1}^{n} w_i m_{ri}, \sum_{i=1}^{n} w_i s_{li}, \sum_{i=1}^{n} w_i s_{ri} \right)_{LR}$ with the left and right reference functions *L* and *R*.

Note that although $\pi_{X_i}, i = 1, 2, ..., n,$ have the same left reference function L and the same right reference function *R*, they may still have different possibility distributions due to different m_{li}, m_{ri}, s_{li}, and s_{ri}.

Now if X_i, $i = 1, 2, .., n$, are n possibilistic loss variables with *LR*-possibility distributions $\pi_{X_i} = (m_{li}, m_{ri}, t_{li}, t_{ri})_{LR}$ of the same left reference function L and the same right reference function *R*, then the aggregate loss $X = \sum_{i=1}^{n} X_i$ has the following properties:

(1) X has *LR*-possibility distribution $\pi_X = \left(\sum_{i=1}^{n} w_i m_{li}, \sum_{i=1}^{n} w_i m_{ri}, \sum_{i=1}^{n} w_i t_{li}, \sum_{i=1}^{n} w_i t_{ri} \right)_{LR}$.

(2) $\bar{M}(\pi_X) = \sum_{i=1}^{n} \bar{M}(\pi_{X_i})$.

(3) $\bar{V}(\pi_X) = \sum_{i=1}^{n} \bar{V}\left(\pi_{X_i}\right)$.

The possibility distribution of possibilistic aggregate loss variable can be explicitly described by the possibilistic distributions of individual possibilistic loss variables. This is not true in general for the probability distribution of probabilistic aggregate loss variable.

6. Conclusions

This paper investigated modeling insurance loss events and the aggregate claim amounts using possibility theory which has been applied successfully in modeling uncertainties for other practical problems. Our possibility distribution model incorporates actuarial experts' knowledge into estimating the possibility distribution for loss variable. An explicit analytical form of possibility distribution of aggregate loss variable is derived. Using possibility distributions approach to model uncertainties in insurance area is in a very early stage. Further applications of possibility methods in pricing insurance products and risk management are needed.

References

[1] Carlsson, C., and Full'er, R. (2001) “On possibilistic mean value and variance of fuzzy numbers,” *Fuzzy Sets and Systems*, 122, 315–326.

[2] Cizek, P., Hardle, W., and Weron, R. (2005) “Statistical Tools for Finance and Insurance”, Springer.

[3] Dubois, D., and Prade, H. (1980) “Fuzzy Sets and Systems: Theory and Applications,” Academic Press, New York.

[4] Dubois D., and Prade H. (1987) “The mean value of a fuzzy number,” *Fuzzy Sets and Systems*, 24, 279–300.

[5] Dubois D., and Prade H. (1988) “Possibility Theory,” Plenum Press, New York.

[6] Dubois, D. and Prade, H. (2000) “Fundamentals of Fuzzy Sets,” Kluwer Academic, Dordrecht.

[7] Guo, L.J. and Huang, Z. (1996) “A Possibilistic Linear Programming Method for Asset Allocation,” *Journal of Actuarial Practice*, 2, 67-90

[8] Huang, Z. and Guo, L.J. (2007) “Possibilistic modeling for loss distributions,” *Working paper*.

[9] Inuiguchi, M. and Tanino, T. (2000) “Portfolio selection under independent possibilistic information,” Fuzzy Sets and Systems, 115, 83-92.

[10] Majlender, P. and Full’er, R. (2003) “On weighted possibilistic mean and variance of fuzzy numbers,” *Fuzzy Sets and Systems*, 136, 363-374.

[11] Shapiro, A. (2004) “Insurance applications of fuzzy logic: the first 20 years,” 39th Actuarial Research Conference.

[12] Verchera, E., Bermúdeza, J.D. and Segurab, J.V. (2007) “Fuzzy portfolio optimization under downside risk measures,” *Fuzzy Sets and Systems*, 158, 769–782.

[13] Zadeh, L.A. (1965) “Fuzzy Sets,” *Information and Control*, 8, 338-353.

[14] Zadeh, L.A. (1978) “Fuzzy sets as a basis for a theory of possibility,” *Fuzzy Sets and Systems*, 1, 3-28

[15] Zimmermann, H.-J. (2001) “Fuzzy Set Theory-and Its Applications,” Fourth Edition, Kluwer Academic Publishers.