



FORECASTING OF CLAIM BY COMBINING PROBABILITY DISTRIBUTIONS

DOI: 10.17261/Pressacademia.2018.959

JBEF- V.7-ISS.3-2018(10)-p.319-324

Elif Makbule Cekici¹, Hakan Aydoğan², Serkan Eti³

¹Marmara University, Faculty of Business, Department of Quantitive Methods, Istanbul, Turkey.

ecekici@marmara.edu.tr, ORCID: 0000-0002-1603-9896

²Kastamonu University, Faculty of Forestry, Department of Forest Industrial Engineering, Kastamonu, Turkey.

hakanaydogan@kastamonu.edu.tr, ORCID: 0000-0001-9482-9888

³Istanbul Medipol University, Istanbul, Turkey.

seti@medipol.edu.tr, ORCID: 0000-0002-4791-4091

Date Received: July 28, 2018

Date Accepted: September 14, 2018

To cite this document

Cekici, E. M., Aydoğan, H., Eti, S. (2018). Forecasting of claim by combining probability distributions. Journal of Business Economics and Finance (JBEF), V.7(3), p.319-324.

Permament link to this document: <http://doi.org/10.17261/Pressacademia.2018.959>

Copyright: Published by PressAcademia and limited licenced re-use rights only.

ABSTRACT

Purpose- The purpose of this study is to focus on forecasting the total claim distribution of fire-disaster and transportation insurances by combining probabilities of claim distribution using the data for 2014 and 2015.

Methodology- In this study, the combination of Weibull and Gamma distributions of fire-disaster and transport policies are discussed. By combining two different probability density function (PDF), a single PDF is derived and thus it is aimed to provide convenience in calculations.

Findings- Two probability distributions were combined with mathematical methods. As a result, a single probability density function was obtained. This probability density function can represent the total claim distribution.

Conclusion- In conclusion, we used the data regarding fire-disaster and transportation insurances for 2014 and 2015. Insurance firms that offer these two policies to their customers need to determine their risk taking into consideration the total claim distribution from the policies. Since Weibull and Gamma distributions have a PDF, it is necessary to integrate the multiplication of the characteristic functions for combining these two distributions. The PDF obtained as a result of mathematical operations is the PDF of the total claim distribution.

Keywords: Claim distributions, combining probability distributions, insurance, total claim distribution, actuarial models.

JEL Codes: G22, C46, C13, E17, C02, C69

1. INTRODUCTION

Businesses and individuals need to take out an insurance policy for daily risk that are unlikely to occur but could cause massive losses if happened. Thus, they aim to reduce the effect from the risks to the minimum level (Inel & Çekici, 2013). In Turkey, there are many insurance policies in the field, especially health. Insurance firms offer policies in many different areas such as health, life, travel, goods, natural disaster, private pension. Through these policies, individuals and firms transfer their risks to insurance companies. Insurance companies make claims payments in case of the realization of the risks they bear against the premiums. Insurance firms need to be able to control their liquidity to minimize the ruin probabilities that may arise from claims payments. For this reason, it is necessary to forecast the claim payments stemmed from the risks that the premiums occurred. Claim payments for these risks depend on the severity and the amount of claim. Moreover, the severity and the amount of claim depend on the case defined and have a mathematical probability distribution.

Since the claim has a probability, it may be possible to forecast the claim payments using probability density function (PDF). Thus, insurance companies can plan their liquidity by forecasting the total claim payment. However, the PDFs of independent claim payments may be different from each other. This makes it difficult to obtain total claim distribution. Hence, many methods have been proposed in the literature to overcome this issue. In this paper, it is aimed to forecast the total claim distribution of fire-disaster and transportation insurances by combining probabilities of claim distributions using the data for 2014 and 2015. For this purpose, we shall discuss claim distributions. We shall then show that methods of combining claim distributions in the literature. Finally, we present an application of the combining of Weibull and Gamma PDFs which correspond to the claim distributions. The probability distribution of claim occurring at a specified time interval is called the claim distribution. Furthermore, it has a feature of continuous random variable and especially, the claim distribution of operational risk losses often shows a positive skewness and high kurtosis. When an appropriate distribution for any random variable is considered, the heavy-tailed distributions are the distributions commonly used due to the nature of the financial data. Moreover, it has been seen in the literature that some commonly used distributions are as follows: Normal, Lognormal, Exponential, Pareto, Burr, Weibull and Gamma distributions.

Risk statistics require urgency when they have been done against time, and it is not possible to complete until all losses have been recognized and/or all claims have been paid. In some fields, as the incurred but unreported claims, actuaries may need to forecast past losses. In particular, it may seem insignificant as this issue can easily be observed in fire insurance; but the payment for large claims can be considerably delayed.

For a future time interval, if S denotes the sum of claims and n denotes the number of claims, then it may be expressed as:

$$S = n \cdot \frac{S}{n}$$

When all information about the policy is taken into consideration, the distribution function $P(y)$ of the individual claim needs to be forecasted to obtain the best information about the S/n . Not only for the application of the collective risk theory, but also for daily decisions with regards to deductibles, initial losses, retentions, loadings and other reinsurance issues, the $P(y)$ is a fundamental function. Since these decisions depend on the tails of the distribution, forecasting is based on as large statistics as possible is necessary. Therefore, the forecasting of $P(y)$ is quite appropriate for the cooperation of competing firms (Benckert and Jung, 1974).

At any time, the sum of the payments made for claim or losses in the insurance company's portfolio is regarded as the total amount of claim. The distribution of the total amount of claim is theoretically calculated by multiplying the distributions of the number of claim and the amount of claim. Moreover, different methods have been developed in the literature to calculate the distribution of the total claim amount.

2. LITERATURE REVIEW

Studies on the combining of probability density functions have not existed much due to the mathematical difficulty. It is seen that the studies are frequently evaluated in a theoretical framework rather than in practice. There are different approaches in the literature on combining probability distributions. While some of these suggest combining characteristic functions of PDFs, the others suggest the Bayesian approach. Apart from these, it is used few methods such as the way of assigning importance to the distributions with the expert opinion or weighting method providing that sum of the distributions is one. The combination of probability distributions is used in many sectors, primarily in the field of finance and insurance. In determining of risks, the combining processes of probability distribution are applied in radar systems as well.

Winkler (1981) has stated that uncertainty may be represented officially with regard to probability, and for the uncertain amount of interest, the crucial purpose of a consensus procedure is to reach at a probability distribution. Additionally, he has claimed that the consensus distribution entirely shows the information that is provided by the experts. Ranging from a simple average of probability distributions to the transition to Bayesian as common ideas of scholars such as Eisenberg and Gale (1959), Stone (1961), Roberts (1965), Winkler (1965), and Morris (1974). While Winkler proposed a consensus model for this process, some experts suggest that a Bayesian approach is a more appropriate choice. Genest and Zidek, in their study, benefited from a Bayesian approach as other academics. Winkler and Clemen, on risk analysis combined the density functions of the risks with experts' opinions. Thus, they determined the risk analysis more precisely. An important feature of information obtained by different sources is the probability of stochastic dependency. The developed model expresses the dependency between faulty estimations of experts.

Expert opinions can be represented by PDFs. Hence, a PDF can be reduced to the combined distribution or consensus distribution. Particularly, when the information is inadequate, the process of decision making can be risky. In order to solve this issue, Liu et al. (2012) presented a proposal. (1) the problem is formulated as an optimization problem, and (2) the overlapping fields of the PDFs are combined with an expert opinion and the level of the expert evaluation is then used to determine the weight of the stated opinion. On the other hand, Lui et al. (2012) presented a novel approach that is about combining the set of PDFs or the distributions. Their approaches are based on the idea that the weight of a decision maker is related to the overlapping field between the PDFs of decision makers and the combined PDF. The combined PDF is represented by a weighted average model. Thus, determining of a combined PDF is reduced to determining the weight for each expert.

In a study on combining the distributions of independent data sets for the same quantity by Hill and Miller (2010), they have first discussed the common method of averaging the probabilities (For each data set of A: $P(A) = (P_1(A) + P_2(A))/2$), however this method has many important disadvantages. Instead, as an alternative to averaging the data or averaging the probabilities, Hill (2008) announced a mathematical method called "conflation". McMillan and Kohlberg (2017) has studied the combining PDFs regarding radar and communication systems, and on the combining statistically well-defined distributions of system problem that are independent of each other in these systems. In their paper, they use PDF calculated for a sum of normal and exponential distributions by using a characteristic function. However, they have claimed that it is quite difficult to combine these distributions using numerical techniques. In fact, it has been stated that in the case of normal and lognormal in the paper, it is even more difficult to combine due to the characteristic function of the lognormal distribution. Hence, they have used the convolution or "faulting" method to combine normal and lognormal distributions as an even simpler method, which had previously been used (McMillan and Barnes, 1976) instead of the characteristic function approach.

Taken together, these studies support the notion that several methods such as the considering of expert opinions, Bayesian approach, convolution, averaging of probabilities, and Conflation have been used to combine different probability distributions.

3. DATA AND METHODOLOGY

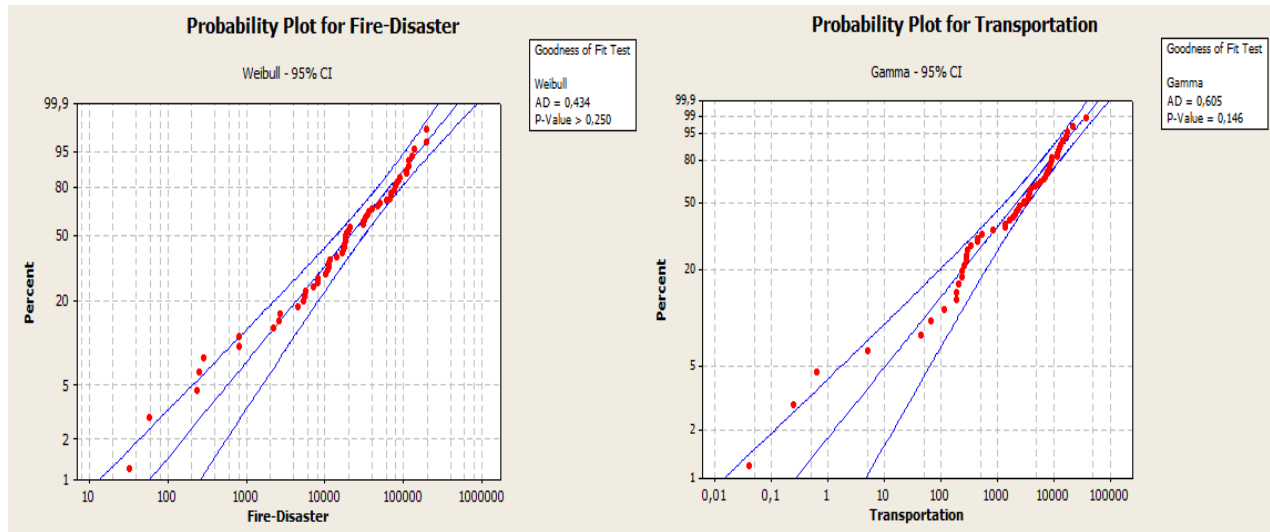
The claim information was obtained by reports at the official website of Republic of Turkey, Ministry of Treasury and Finance. The claim premiums paid for fire-disaster and transportation policies for 2014 and 2015 are given Table 1.

Table 3: Claim premiums paid for Fire-Disaster and Transportation Policies

2015		2014		2015		2014	
Fire-Disaster	Transportation	Fire-Disaster	Transportation	Fire-Disaster	Transportation	Fire-Disaster	Transportation
5656,81	3036,72	799,48	444,78	34480,3	1405,85	77341,79	3792,76
18993,41	6859,67	8194,6	11655,04	32628,8	22240,5	30281,21	206,46
109488,31	9396,95	86228,42	6975,92	232,46	0,64	36031,04	12842,47
90323,69	14373,94	114471,14	36873,28	18083,67	296,4	245,89	0,04
193843,87	13024,54	109373,53	17487,65	5433,67	283,77	20225,27	113,37
18191,75	191,6	17535,24	285,28	79597,49	8256,37	5365,54	443,28
195270,13	7772,43	11665,2	3504,13	16469,63	67,73	60348,16	7411,8
284,76	0,24	128688,65	9177,93	2545,6	840,74	13945,65	234,49
58	253,05	32,46	5,12	80525,08	1628,65	32074,64	1913,55
11252,53	3482,75	10135,18	2181,93	8187,46	3674,58	7104,08	537,55
17647,57	16326,62	799,51	189,35	67885,9	3983,42	30505,6	2928,47
70153,15	8455,59	10661,09	11510,46	11014,34	2353,41	2666,03	336,55
4492,5	5209,33	50505,17	4630,69	18015,52	2443,49	20456,16	44,31
114765,61	5868,3	2196,39	1380,38	39738,83	232,67	17334,73	11886,24
69550,85	2085,07	135976,92	7862,15	46300,18	8970,86		

The distribution of the data for these two policies were examined and tested using Minitab 16 software. We found that the claim distribution of fire-disaster insurance fits to the Weibull distribution, while the claim distribution of transportation insurance fits to the Gamma distribution. Anderson test Darling test was used as a conformity test. According to the results of this test, distributions where claim distributions are appropriate are determined.

Figure 1: The Distribution Test Charts of Fire-Disaster and Transportation Policies



4. FINDINGS AND DISCUSSIONS

In accordance with the theory, Weibull and Gamma distributions were combined. For the functions to be combined, X is the continuous random variable and has the PDF with the $f_X(x|\lambda)$ lambda parameter, and the second density function to be combined is with the $f_\Lambda(\lambda|\theta)$, θ parameter, the combined distribution function can be calculated as follows:

$$f_Y(y|\theta) = \int_{\lambda \in \Omega_\Lambda} f_X(y|\lambda) f_\Lambda(\lambda|\theta) d\lambda$$

Where Y is combined distribution with the parameter of θ . This distribution has a continuous distribution.

With the help of the above-mentioned theory, A.D. the distributions obtained from the test are combined. The density functions of the distributions that are appropriate for the given data are as follows:

$$\text{Weibull } f(y; \gamma, \beta) = \frac{\gamma}{\beta} y^{\gamma-1} e^{-\frac{y^\gamma}{\beta}} \quad 0 < y < \infty, \gamma > 0, \beta > 0$$

$$\text{Gamma } f(x) = \frac{1}{\beta^\alpha \Gamma(\alpha)} x^{\alpha-1} e^{-x/\beta} \quad x > 0, \alpha > 0, \beta > 0$$

The combination of these two PDFs,

$$\iint \left(\frac{\gamma}{\beta} y^{\gamma-1} e^{-\frac{y^\gamma}{\beta}} \right) \left(\frac{1}{\beta^\alpha \Gamma(\alpha)} x^{\alpha-1} e^{-x/\beta} \right) dx dy = \int \frac{\gamma}{\beta} y^{\gamma-1} e^{-\frac{y^\gamma}{\beta}} dy \int \frac{1}{\beta^\alpha \Gamma(\alpha)} x^{\alpha-1} e^{-x/\beta} dx$$

is obtained by solving the integral. For the integral solution, when the $-\frac{y^\gamma}{\beta} = t$ and $-\frac{x}{\beta} = u$ are transformed, the combination of the two density functions is given by

$$\gamma e^{-\frac{y^\gamma}{\beta}} \frac{1}{(n-1)!} e^{-\frac{x}{\beta}} \left(-\frac{x}{\beta} - 1 \right)$$

This solution is the probability density function of the total claim distribution, which is the combination of two claim distributions. By estimating of the parameters involved in this combination will lead to forecast of the total claim. The claim distribution is reduced to a single probability distribution by the PDF. Systematically, using of multiple PDF increases the probability of Type 1 error. It also increases the degree of freedom. Hence, using a single reduced PDF will enable more accurate.

5. CONCLUSION

Firms and individuals require to take out a policy for daily risk that are unlikely to occur but might cause high losses if happened. Insurance firms that transfer their risks should correctly assess the payment probabilities of the losses stemming from these risks and must overcome the liquidity issue. Thus, it is crucial for insurance companies to determine distributions of claim payments correctly and also to express the total claim in a single distribution. Handling of individual claim distributions separately will increase confidence interval as well as cost and degrees of freedom. Achieving the probability function of the total claim distribution in order to forecast a more precise claim payment is an extremely important matter. With this consideration, it is necessary to combine two probability functions mathematically. Much of the existing literature on combining probability distributions pays particular attention to each claim distribution is different from each other. Since the claim distributions are different, it is difficult to calculate the total claim distribution. The purpose of the current study was to provide great convenience to insurance companies about forecasting of their total claim distribution by combining probabilities of claim distributions. For this purpose, we used the data regarding fire-disaster and transportation insurances for 2014 and 2015. Insurance firms that offer these two policies to their customers need to determine their risk taking into consideration the total claim distribution from the policies. Thus, it was attempted to combine the PDFs of the two policies. While the distribution of claim payments for fire-disaster policies fitted the Weibull distribution, the distribution of claim payments for transportation policies was determined to be appropriate for Gamma distribution.

Theoretically, since Weibull and Gamma distributions have a PDF, it is necessary to integrate the multiplication of the characteristic functions for combining these two distributions. The PDF obtained as a result of mathematical operations is the PDF of the total claim distribution. Apart from this study, the combining of normal and lognormal distributions is also discussed in the literature. As the claim distributions can have many distributions besides the distributions discussed, it is recommended that combining of other distributions should be considered and evaluated in the future studies.

REFERENCES

- Armutlu, İ. H. (1999). İşletme istatistiğine giriş. Beta Basım Yayım Dağıtım.
- Benckert, L. G., Jung, J. (1974). Statistical models of claim distributions in fire insurance. ASTIN Bulletin: The Journal of the IAA, 8(1), 1-25.
- Clemen, R. T., Winkler, R. L. (1999). Combining probability distributions from experts in risk analysis. Risk analysis, 19(2), 187-203.
- Çekici, M. E., İnel, M. N. (2013). Türk sigorta sektörünün direkt prim üretimlerinin tahmin teknikleri ile incelenmesi. Marmara Üniversitesi İ.İ.B. Dergisi, 34(1), 135-152.
- Eisekberg, E., Gale, D. (1959). Consensus of subjective probabilities: the pari-mutuel method. Ann. Math. Statist. Vol. 30 (1959), pp. 165-168.
- Genest, C., Zidek, J. V. (1986). Combining probability distributions: a critique and an annotated bibliography. Statistical Science, 1(1), 114-135.
- Hill, T. P., Miller, J. (2011). How to combine independent data sets for the same quantity. Chaos: An Interdisciplinary Journal of Nonlinear Science, 21.3 (2011): 033102-1.
- Hill, T. (2011). Conflations of probability distributions. Transactions of the American Mathematical Society 363.6 (2011): 3351-3372.
- Liu, X., Ghorpade, A., Tu, Y. L., Zhang, W. J. (2012). A novel approach to probability distribution aggregation. Information Sciences, 188 (2012), 269-275.
- McMillan, R. W., Barnes, N. P. (1976). Detection of optical pulses: the effect of atmospheric scintillation. Applied Optics, Vol. 15, No. 10, Ekim 1976.
- McMillan, R. W., Kohlberg, I. (2017). A simple method for combining probability distribution functions relevant to radar and communications systems. IEEE International Conference on Microwaves, Antennas, Communications and Electronic Systems (COMCAS), (2017).

Morris, P. A. (1974). Decision analysis expert use. *Management Sci.*, Vol. 20 (1974), pp. 1233-1241.

Pırıldak, M., Esensoy, Ö. Bağımlı riskler için toplam hasar miktarının dağılımı. *İstatistikçiler Dergisi: İstatistik ve Aktüerya*, 1(1), 68-79.

Roberts, H. V. (1965). Probabilistic prediction. *J. Amer. Statist. Assoc.*, Vol. 60 (1965), pp. 50-62.

Stone, M. (1961). The opinion pool. *Ann. Math. Statist.*, Vol. 32 (1961), 1339-1342.

Tse, Y. K. (2009). *Nonlife actuarial models: theory, methods and evaluation*. Cambridge University Press.

Winkler, R. L. (1968). The consensus of subjective probability distributions. *Management Sci.*, Vol. 15 (1968), pp. B61-B75.

www.hazine.gov.tr