MTPL RATES: A STATISTICAL MODEL FOR THE ALBANIAN INSURERS

[1]Eralda Dhamo (Gjika), [2]Oriana Zaçaj, [3]Miftar Ramosaço, [4]Klodiana Bani

[1,4]Department of Applied Mathematics, Faculty of Natural Science, University of Tirana, Albania

[2]Department of Mathematical Engineering, Faculty of Mathematical Engineering and Physical Engineering, Polytechnic University of Tirana, Albania

[4]Department of Mathematics, Faculty of Technical Sciences, University of Vlora, Albania

e-mail: eralda.dhamo@fshn.edu.al

Abstract: In some of the countries where this system is implemented (including Albania) the motor third party liability constitutes the main insurance portfolio – in Albania in 2014 the compulsory motor insurance share was 65% of the total premium income and more than 70% of the claims paid (AMF - Insurance Market Statistical Report).
Therefore, while it is generally important for the MTPL premium rate to be fairly stipulated to guarantee that the insurer shall always be able to indemnify the injured/damaged party, this becomes even more important in countries like Albania if we only consider that potential issues in the motor portfolio directly influence the market stability because of the relatively high share of the this type of insurance in the overall insurance market.

In this paper we will find a probability model that better fits to the claim data for the MTPL portfolio of an Albanian insurance company. Then we will use the model to calculate the premium rates for that company, which may serve as a model for the Albanian Insurance Market

Key words: *claims, probability, models, premiums, R.*

1. Introduction

MTPL (Motor Third Party Liability) insurance is the insurance of the liability of the owner/user of the motor vehicles against losses/damages caused to third parties resulting from the use of the said motor vehicle.

Taking into consideration the high level of risk in operating a motor vehicle, the governments establish by law and implement financial protection systems to financially compensate the damaged party for bodily injuries and/or material damage resulting from a road accident.

Of special importance is the financial protection of victims from road accidents (resulting in partial bodily injuries or loss of life). The WHO (World Health Organisation) data show that 1.2 million people die and more than 20 million more sustain various bodily injuries as a result of road accidents. The financial protection systems provide for the compulsory third party liability insurance of the owner/user of the motor vehicle, to ensure that, despite the financial means of the party liable for the accident, the injured/damaged party receives the due indemnity from the insurance company (in case that the owner/user of the motor vehicle, in violation of the law, is uninsured, the indemnity is paid out of a fund specially established for this purpose).

In some of the countries where this system is implemented (including Albania) the motor third party liability constitutes the main insurance portfolio - in Albania in 2014 the compulsory motor insurance share was 65% of the total premium income and more than 70% of the paidclaims (AMF - Insurance Market Statistical Report).

Therefore, while it is generally important for the MTPL premium rate to be fairly stipulated to guarantee that the insurer shall always be able to indemnify the injured/damaged party, this becomes even more important in countries like Albania if we only consider that potential issues in the motor portfolio directly influence the market stability because of the relatively high share of the this type of insurance in the overall insurance market.

2. Risk premium

The rating process normally starts with a calculation of the pure risk premium (i.e. the premium required simply to meet the expected cost of claims arising from the policies written under the new rates); to this we then have to add loadings for expenses, profit and other contingencies.[5]

We determine the risk premium from the base data by deriving appropriate values for exposure and claims. We then project there values to the mid-point of the new exposure period covering the policies written under the new rates.

We should make allowance in the projection of these values for inflation and any trends in experience thought likely to apply over this period.

There are Many factors affect risk premium such as number of claims, claim cost, number of earned contracts etc.

The premium calculation formula is:

$$
\begin{aligned}
Risk\ Premium &= \frac{Cost}{Exposure} = \\
&= \frac{Number\ of\ Claims}{Number\ of\ contracts} * \frac{Claim\ Cost}{Number\ of\ Claims} * \frac{Exposure\ units}{Number\ of\ contracts} = \\
&= \frac{Number\ of\ Claims}{Number\ of\ earned\ contracts} * \frac{Claim\ Cost}{Number\ of\ Claims} = \\
&= \frac{Claim\ Frequency}{Number\ of\ earned\ contracts} * Claim\ Severity
\end{aligned}
$$

In this paper we'll try to calculate each of these components for the MTPL portfolio of an Albanian insurer.

2.1. Claim frequency distribution

An important measure of claim losses is the claim frequency, which is the number of claims in a block of insurance policies over a period of time [6], [7], [8]. Though claim frequency does not directly show the monetary losses of insurance claims, it is an important variable in modeling the losses. The claim freequency can be modeled as a non negative discrete distribution. The most used distribution may be:

Binomial: $f(x,n,p) = P(X=x) = C_n^x p^x q^{n-x}, x = 0,1,2,...,n$

Poisson: $f(x,a) = P(X=x) = \dfrac{e^{-a}a^x}{x!}, x = 0,1,2,...$

Negative Binomial: $f(x,r,p) = P(X=x) = C_{r+x-1}^x p^x q^r, x = 0,1,2,...$

Geometric: $f(x,p) = P(X=x) = (1-p)^{x-1}p, x = 1,2,...$

Apart from these distributions, we could use techniques like compound distributions, and the mixture distributions for creating new distributions which may fit to the claim frequency data.

2.2. Claim severity distribution

The aggregate claims for losses of the block of policies, is the sum of the monetary losses of all the claims [6], [7], [8]. Unlike claim frequency, which is a nonnegative integer-valued random variable, claim severity is usually modeled as a nonnegative continuous random variable. The most used distribultion may be:

Exponential: $f(x,\lambda) = \lambda e^{-\lambda x}, x > 0$

Gamma: $f(x,k,\theta) = \dfrac{x^{k-1}e^{-\frac{x}{\theta}}}{\theta^k\Gamma(k)}, x > 0, k > 0, \theta > 0$

Weibull: $f(x,\lambda,k)=\frac{k}{\lambda}(\frac{x}{\lambda})^{k-1}e^{-(\frac{x}{\lambda})^k}, x>0$

Lognormal: $f(x)=\frac{1}{x\sigma\sqrt{2\pi}}e^{-\frac{(\ln x-\mu)^2}{2\sigma^2}}, x>0$

Pareto: $f(x,\alpha)=\frac{\alpha x_m^{\alpha}}{x^{\alpha+1}}, x>x_m, \alpha>0$

Apart from these distributions, we could use teqniques such as the mixture-distribution method for creating new distributions which may fit to the claim severity data. Claims-size models are often used to support pricing high-level insurance coverage. This type of coverage tends to have somewhat “thick” tails due to a few huge claims having a disproportional effect on the average.

3. MTPL Portfolio of an Albanian insures

3.1. Claim Frequency distribution

Claim Frequency probability distribution was choose between Poisson and Negative binomial. Based on previous researchers in foreign countries.[1], [6], [7], [8].

After performing graphical test, and goodness of fit tests for accuracy of probability distribution we conclude that Negative Binomial model is a better approximation of Frequency. [1], [3], [4]

Table 1. Summary of estimation results for claim frequency

Distribution	λ	
Poisson	119.0526	387.47
	r	θ
Negative binomial	3.68	0.0277

Table 2: The resulting statistics for Chi Square test

	Poisson	**Negative Binomial**
Statistic	4051.05	**11.41609**
Degrees of freedom	10	**9**

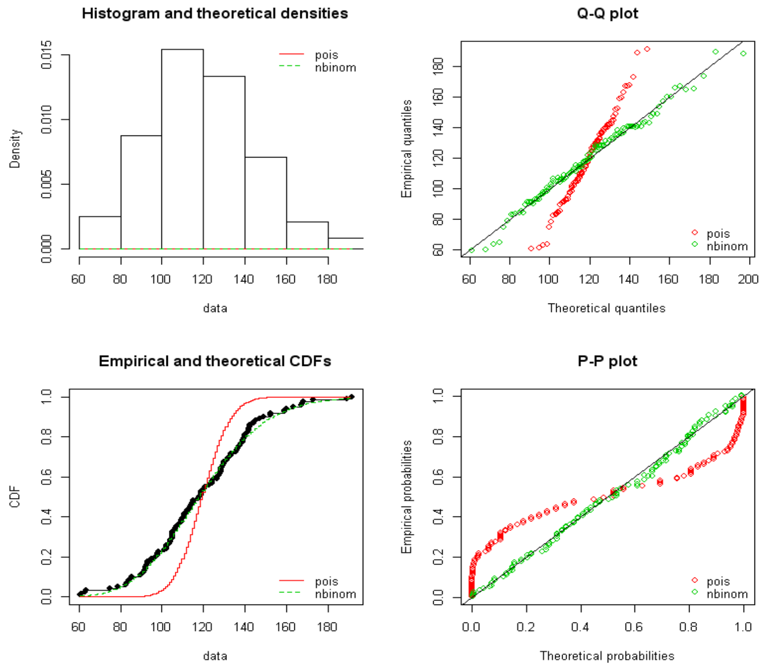

Figure 1. Histogram,Q-Q plot, Empirical and theoretical CDFP-P plot for the claim frequency distribution

3.2. Claim Severity distribution

Claim Severity probability distribution was choose between Weibull and Lognormal. Based on previous researchers in foreign countries. [1], [2], [6], [7], [8].

After performing graphical test, goodness of fit tests and information criteria for accuracy of probability distribution we conclude that the lognormal model is a better rapproximation of severity distribution.

Model selection

A number of measures contemplate the number of parameters in determining the fit of a model. While none have been universally accepted as an appropriate approach, two types of measurement – the Bayesian Information Criterion (BIC) and the Akaike Information Criterion (AIC) – are important to consider. Both begin with the negative log-likelihood value for a particular fit and modify that value to reflect the sample size, n, and the number of parameters of the model that are fitted, *p*.

We have based our "best model selection" procedure on Akaike Information criterion (AIC, AICc), Bayesian Information criterion (BIC), and other statistical test such as: *Kolmogorov-Smirnov statistic, Cramer-von Mises statistic, Anderson-Darling statistic.*

Akaike Information Criterion: $AIC = -2\log(Likelihood) + 2p$

where p is the number of estimated parameters in the model. Minimizing the AIC gives the best model. Because AIC does not have much meaning by itself so it is useful in comparison to AIC value for another model fitted to same data set. So we compare the AIC value for many probability distribution of our data. [1], [2], [3], [4],[6], [7], [8], [9]

Schwartz' Bayesian: $BIC = AIC + p(\log(n) - 2)$

Table 3. Summary of estimation results for claim severity

Distribution	Shape	Scale	AIC	BIC	Log likelihood
Weibull	2.413	1.57	2,974.721	2,980.296	-1,485.36
	Meanlog	**Sdlog**	**AIC**	**BIC**	**Loglikelihood**
Lognormal	11.755	0.407	**2,950.179**	**2,955.754**	**-1,473.09**

The lowest value of AIC and/or BIC gives the best probability distribution model for our data. So, as can be observed by Table 3 Lognormal distribution better fits the claim severity probability distribution.

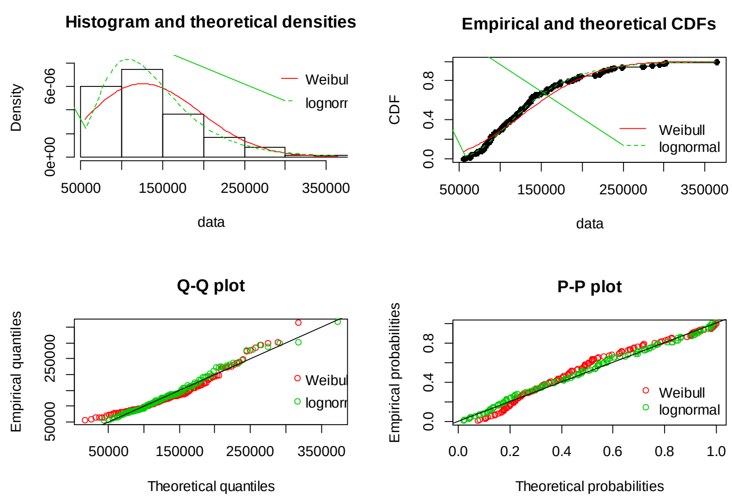

Figure 2. Histogram,Q-Q plot, Empirical and theoretical CDFP-P plot for the claim severity distribution

Figure 2 show some of the graphical test we have used to compare the performance and godness of fit of the selected probability distributions. In all graphical tests the Lognormal probability distribution show to better fit the claim severity.

To ensure that the best probability distribution function is Lognormal we have performed three other statistical tests: *Kolmogorov-Smirnov statistic, Cramer-von Mises statistic, Anderson-Darling statistic*. The model which have the lower value of these statistical test is the "best" probability model to fit the claim severity distribution.A powerful goodness-of-fit test, the Kolmogorov-Smirnov is more often used for testing uniformity of the proposed set of random numbers. This is done by conducting tests on the greatest deviation between the empirical distribution of the random numbers and the uniform distribution[1], [2], [3], [4].

Table 4. The resulting statistics for some tests

	Weibull	**Lognormal**
Kolmogorov-Smirnov statistic	0.1142	**0.0548**
Cramer-von Mises statistic	0.3682	**0.0679**
Anderson-Darling statistic	2.3806	**0.4761**

3.3. The premium calculation

Based on the results from claim frequency and claim severity above, the components for the risk premium of the MTPL portfolio of the albanian insurer considered in our paper are: [1], [2], [3], [4], [5], [6], [7], [8].

Table 5. The resulting components of the risk premiumof the MTPL portfolio of the albanian insurer

	Mean
Claim frequency (Negative binomial)	1,428.6312
Claim Severity (lognormal)	138,389.09
Earned Contracts	38,553

And the risk premium becomes initially

$$Risk\_\Pr emium = \frac{Claim\_Frequency}{Number\_of\_earned\_contracts} * Claim\_Severity$$

$$= \frac{1429}{37553} * 138389 = 5128.19$$

If we consider a safety component of 15% for claim frequency, 50% for Claim Severity and 15% for inflation and Compensation Fund, we have:

$$Risk\_\Pr emium = 10173.04$$

For the **office premium** we have to consider other components such as administrative expenses and profit. If we consider these expense loadings as approximately 25% of the risk premium, then we have:

$$\begin{aligned} Office\_\Pr emium &= risk\_\Pr emium + (1 + Expenses) \\ &= 10173.04 * 1.25 \\ &= 12716.3 \end{aligned}$$

4. Conclusions:

We found a probability model that better fits to the claim frequency distribution and claim severity data for the MTPL portfolio of an Albanian Insurance Company.

Negative Binomial and Lognormal distributions were found to be more adaptive for the data. (Claim Frequency and Severity)

Using these probability distributions we calculated the premium rates for that portfolio

From the claim experience of the company in our consideration we find that even the average premium in the Albanian Market results approximately ALL 16,286, in our calculations it results not greater than ALL 12,716

References

[1] Zacaj O., Dhamo E., (2015): Trial Modelling For MTPL Claims In Albania, Journal of Multidisciplinary Engineering Science and Technology (JMEST) Vol. 2 Issue 6, June - 2015 (ISSN: 3159-0040, pg 1479)

[2] Zacaj O., Dhamo E., Shehu Sh. (2015): Modeling Individual Claims for Motor Third Party Liability of Insurance Companies in Albania International Journal of Innovation in Science and Mathematics Volume 3, Issue 3,May 2015 (ISSN (Online): 2347–9051, pg 174)

[3] Delignette-Muller, M., Pouillot, R., Denis, J., and Dutang, C. (2014). fitdistrplus: Help to Fit of a Parametric Distribution to Non-Censored or Censored Data. R package version 1.0-2.

[4] Kohl, M. and Ruckdeschel, P. (2010). R Package distrMod: S4 Classes and Methods for Probability Models. Journal of Statistical Software.

[5] Benjamin, B. (1991). General Insurance. Institute and the Faculty of Actuaries

[6] S. A. Klugman, H. H. Panjer, G E. Willmot,(2004) Loss Models From Data to Decisions,2nd edition

[7] YIU-KUEN TSE, (2009), Nonlife Actuarial Models Theory, Methods and Evaluation

[8] Roger J. Gray, Susan M. Pitts "Risk modeling in general insurance: From Principles to Practice

[9] Hyndman R. J., Koehler, A. B., Snyder, R. D. & Grose, S. (2002): A state space framework for automatic forecasting using exponential smoothing methods, International Journal of Forecasting 18, 439–454]