PLISKA
STUDIA MATHEMATICA

ПЛИСКА
МАТЕМАТИЧЕСКИ
СТУДИИ

Pliska Stud. Math. **25** (2015), 147–154

PLISKA
STUDIA MATHEMATICA

MODELING ISSUES OF THE CLAIM PROCESS AND INSURANCE RISK

Elitsa Raeva, Velizar Pavlov

This work presents a brief overview of some proper statistical distributions for modeling of claims and risk in general insurance. Applications of normal approximation, normal power transformation and power transformation for modeling the number and the size of claims have been considered. Some problems in terms of modern risk theory are described and their reformulation in the actuarial practice are made. There are numerical examples to determine the amount of insurance premiums in order to limit to one percent the probability of insolvency. Some comparisons were made.

1. Introduction

The modeling of the Claim process in General Insurance is based on estimation of the probability distribution of the individual claims size which occur, as well as their expected number and frequency during some given moment. These variables could be considered as three single stochastic processes, whose characteristics are independent of each other. Combining these three independent processes makes possible the composing of a complex process, which gives information about the distribution of the total number and total value of the claims. The insurance risk, in turn, is defined as the standard deviation of the distribution of the total size of the claims. Thus, the choice of the method of estimation of the claims process is one of the basic stages when modeling the risk.

2010 *Mathematics Subject Classification:* 62C42, 62G22.
Key words: statistical distribution, models, insurance, claim, risk.

2. Number of claims

To be possible the consideration of the number of the claims as a stochastic process, the following conditions should be fulfilled:

- The claims should occur independently of each other;
- Only one claim could be required at a given moment.

The distribution, which is most often connected with the counting as a random process is the Poisson distribution. The main advantages are expressed in the fact that there is only one parameter for estimating and it is additively.

Another probability distribution, which is used for modeling the number of claims is the negative binomial distribution. Let given random variable X has Poisson distribution with parameter q. $X \sim Po(q)$. If the parameter q is random variable with Gamma distribution and parameters r and $\theta = \dfrac{p}{1-p}$, then with some transformations could be shown that the random variable X actually has Negative binomial distribution with parameters r and p, i.e. $X \sim BN(r, p)$. This relation leads the additive property and makes the negative binomial distribution also widely used, [6].

If the expected number of claims is large enough (more than 50), then the Central Limit Theorem could be applied and the distribution of the number of the claims may be approximated with Normal distribution. With small number of the claims and more homogeneity values of the observations, the Poisson distribution could be prefered, while in the cases of larger deviations of the values of the expectation of the number of the claims, the negative binomial distribution may lead to more accurate results.

3. Size of the individual claims

When considering the size of the individual claims, the most widely used distributions are Log-normal, Gamma distribution and Pareto distribution. The shape of these distributions is characterized with the following: they are one-sided, because negative claims cannot occur; strongly asymmetric; they are concentrated around the mode, [1].

4. Total size of the claim process

Let us consider an insurance portfolio, which includes a number of claims. The mathematical mechanism, which can help to present the total size of claims at the end of some given period, is compounding of probability distributions, [2].

Let's denote with C the total value of the claims. By convolution could be expressed the conditional probability, if during the year have been occurred exactly n claims, then their total value to be exactly C, $f^{n*}(C)$:

$$f^{n*}(C) = f * f^{(n-1)*}(C) = \sum_{X=0}^{C} f(X) f^{(n-1)*}(C-X), \tag{1}$$

where: X is the size of the individual claim;
$f(X)$ is the probability given claim to cost exactly X.

In a simmilar manner with the distribution function $F(X)$ could be presented the probability $F^{n*}(C)$ - if during the year have been occurred n claims, their total value to be less or equal to C.

$$F^{n*}(C) = F * F^{(n-1)*}(C) = \sum_{X=0}^{C} F(X) F^{(n-1)*}(C-X). \tag{2}$$

Thus from (1) and (2), using composition of the distributions of the number and the size of the claims, we achieve an expression for the total value of the claims, to be exactly C or respectively less or equal to C:

$$g(C) = \sum p(n) f^{n*},$$

$$G(C) = \sum p(n) F^{n*},$$

where with $p(n)$ denote the probability to occur exactly n claims during the year.

In many cases, however, is difficult to obtain the expressions above in explicit or applicable form. For this reason, if there are enough observations, usually are applied different approximation methods of the distributions. Let the distribution of the individual claims has finite moments. Using the result from Central Limit Theorem, the total value of the claims could be approximated with normal distribution.

$$G(C) = \Phi(C; n\mu, n\sigma^2), \quad \text{where} \quad \Phi(x;\mu,\sigma^2) = \frac{1}{2\pi\sigma} \int e^{-\frac{(x-\mu)^2}{2\sigma^2}} dx \tag{3}$$

With μ and σ^2 denote the expectation and the dispersion of the distribution of the individual claims respectively.

When there is significant asymmetry ot larger coefficient of variation of the values, the Normal aproximation (3) doesn't give enough accurate result. There are different transformations of the random variables, which include information about the coefficient of asymmetry and/or the dispersion, and in such way they give more proper description of the distribution shape. Such examples are Normal power transformation (4) and Power transformation (5), where the probability functions obtain the following form, respectively

$$G(x) = \Phi\left(\frac{-3}{\gamma} + \sqrt{1 + \frac{6x}{\gamma} + \frac{9}{\gamma^2}}\right) \tag{4}$$

$$G(x) = \begin{cases} \Phi(x^h) & \text{if} \quad h \neq 0, \\ \Phi(\ln x) & \text{if} \quad h = 0, \end{cases} \tag{5}$$

where γ is the coefficient of asymmetry, and $h = 1 - \dfrac{\gamma}{3\sigma^2}$

5. Comparing the models

For comparison of the models, let us suppose, that the total size of the claims has Gamma distribution, $X \sim Gamma(k, \theta)$. Then

$$F(X) = \int_0^x f(u; k, \theta) du = \frac{\gamma(k, \frac{x}{\theta})}{\Gamma(k)}, \text{ for } x > 0, \text{ and } k, \theta > 0,$$

where $\gamma(s, x) = \displaystyle\int_0^x t^{s-1} e^{-t} dt$, $\Gamma(s, x) = \displaystyle\int_x^\infty t^{s-1} e^{-t} dt$.

Let us consider the probability for this, the value of the claims to exceed given value C, $(1 - G(C))$. Define the errors from the approximation of Gamma with Normal distribution in the following way:

$$Error = \left| \frac{1 - G(x)}{1 - \Phi(x)} - 1 \right|.$$

If calculate the result of the above formula for different values of the coefficient of asymmetry, using Normal approximation, Normal power transformation and Power transformation, we obtain the following numerical results for the errors:

From the three tables one can see that when transformation of the random variable is used, the approximation with Normal distribution leads to smaller values of the errors in the considered domain, unlike the direct Normal approximation, where the values are of the order of 10^7, [3]. The huge values of the errors mean that, the denominator $1 - \Phi(x)$, i.e. the probability that the cost of the claims exceeds given value is very close to 0 for some values of γ. Such 'zero risk' could be misleading.

Table 1. Errors in the case of Normal approximation

Skewness	Error						
γ/x	0.5	1.0	1.5	2.0	3.0	4.0	5
0.30	2e-17	3.6e-11	5.9e-11	9.7e-10	2.5e-10	6.5e-10	1.6e-09
0.50	5e-05	8.8e-05	1.4e-04	2.3e-04	5.8e-04	0.001	0.003
0.75	0.007	0.011	0.017	0.024	0.034	0.022	0.013
1.00	0.039	0.051	0.045	0.019	0.064	0.133	0.141
2.00	0.123	0.264	0.277	0.147	1.19	12.57	211.8
3.00	0.398	0.328	0.250	2.9	185.9	790e+02	10e+07

Table 2. Errors in the case of Normal power transformation

Skewness	Error						
γ/x	0.5	1.0	1.5	2.0	3.0	4.0	5
0.30	2.2e-11	3.6e-11	5.6e-11	8.5e-11	1.9e-10	3.9e-10	7.7e-10
0.50	5.6e-05	8.8e-05	1.3e-04	1.9e-04	3.8e-04	6.9e-04	0.001
0.75	0.007	0.011	0.015	0.019	0.009	0.043	0.147
1.00	0.045	0.051	0.030	0.021	0.195	0.398	0.581
2.00	0.031	0.264	0.425	0.540	0.686	0.771	0.824
3.00	0.156	0.328	0.397	0.419	0.377	0.236	0.030

Table 3. Errors in the case of Power transformation for $h = \dfrac{1}{3}$

Skewness	Error						
γ/x	0.5	1.0	1.5	2.0	3.0	4.0	5
0.30	2.9e-11	3.6e-11	4.1e-11	4.6e-11	5.6e-11	6.4e-11	7.3e-11
0.50	7.2e-05	8.8e-05	1.0e-04	1.1e-04	1.4e-04	1.5e-04	1.1e-04
0.75	0.009	0.011	0.012	0.010	0.013	0.085	0.207
1.00	0.056	0.051	0.012	0.063	0.281	0.511	0.697
2.00	0.043	0.264	0.496	0.660	0.849	0.934	0.972
3.00	0.063	0.328	0.517	0.654	0.824	0.912	0.956

6. Insurance risk

The insurance risk is defined like the possibility to occur a event, which require a payment of a claim from the side of the insurer. In practice, the purpose is to estimate the risk of loss and to minimize this probability of loss.

Consider an insurance portfolio with free reserves U and safety loadings λ. The number of the claims is random variable N with expectation $E\{N\} = n$. With $X_i, i = 1, 2, \ldots, N$ denote the cost of the i-th claim, for which $E\{X_i\} = m$, $E\{X_i^2\} = \alpha_2$. In the descrete case, the total value of the claims is presented as a sum of a random number of random variable as follows:

$$C = \sum_{i=1}^{N} X_i,\ E\{C\} = nm.$$

The probability for insolvency could be defined as

$$Pr\{C > U + (1+\lambda)nm\}.$$

Let ϵ is a given maximal permissible level of bankruptcy. Then we can write the following condition

$$Pr\{C > U + (1+\lambda)nm\} \leq \epsilon.$$

After standardizing the random variable C, the above inequality remains in force and transforms in

$$Pr\{\frac{C-nm}{\sqrt{n\alpha_2}} > \frac{U+\lambda nm}{\sqrt{n\alpha_2}}\} \leq \epsilon \tag{6}$$

Thus, in a sufficient number of the claims from Central Limit Theorem follows that, $\frac{C-nm}{\sqrt{n\alpha_2}} \sim N(0,1)$. If denote $\Phi^{-1}(1-\epsilon) = z_\epsilon$, then (6) can be transformed in

$$\frac{U+\lambda nm}{\sqrt{n\alpha_2}} \geq z_\epsilon.$$

From here we obtain constraint for the free reserves, depending of the distribution of the number of the claims, the size of the individual claims, the maximum permited level ϵ and the safety loadings, [4].

$$U \geq z_\epsilon\sqrt{n\alpha_2} - \lambda nm \tag{7}$$

Similarly reasoning is valid also for the safety loadings.

$$\lambda \geq \frac{z_\epsilon\sqrt{n\alpha_2} - U}{nm}$$

If in this case is applied, for example, Normal power transformation, then

$$U = (z_\epsilon + \frac{\gamma(z_\epsilon^2-1)}{6})\sqrt{n\alpha_2} - \lambda nm \tag{8}$$

From the last formula it can be seen that, when the ciefficient of asymmetry has positive value, the safety loadings are increasing proportional to the increasing of γ. Consequently, in a case of positive asymmetry in the claims distribution, the Normal power transformation leads to higher lower bound for U and λ respectively.

Another way for reducing the probability of insolvency is through to reinsurance of a given sum from the insurance portfolio, called retention M. Let the distribution of the size of the claims to be continuous, with probability density function $s(x)$. Then, for the second moment of the size of the individual claims could be obtained the following estimation

$$\begin{aligned} \alpha_2 &= \int_0^M x^2 s(x)dx + M^2\int_M^\infty s(x)dx \\ &\leq M\int_0^M xs(x)dx + M^2\int_M^\infty s(x)dx \\ &\leq M\left(\int_0^M xs(x) + M\int_M^\infty s(x)dx\right) \leq Mm, \end{aligned}$$

where m is the expectation of the distribution of the individual claim. Let $k \leq 1$ is coefficient, which presents factor, depending of the distribution of the claim.

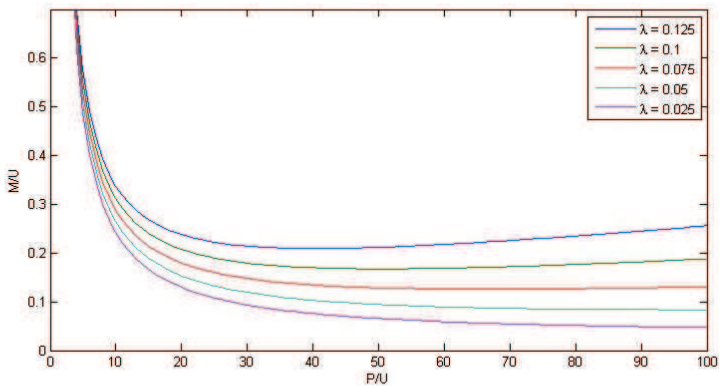

Figure 1: Normal Approximation

Then we can write

$$\alpha_2 = k \times Mm$$

If we ignore the safety loadings and accept that, $P = mn$ and lay in (7) (8), then

$$U = z_\epsilon \sqrt{kPM} - \lambda P$$

and

$$U = (z_\epsilon + \frac{\gamma(z_\epsilon^2 - 1)}{6})\sqrt{kPM} - \lambda P,$$

from where

$$M = \begin{cases} \dfrac{(U+\lambda P)^2}{(z_\epsilon^2 kP)}, & \text{Normal approximation} \\ \dfrac{(U+\lambda P)^2}{(z_\epsilon + \frac{\gamma(z_\epsilon^2-1)}{6})^2 kP}, & \text{Normal power transformation.} \end{cases}$$

After transforming the equality above, we can obtain

$$\frac{M}{U} = \begin{cases} \dfrac{(1+\lambda \frac{P}{U})^2}{(z_\epsilon^2 k \frac{P}{U})}, & \text{Normal approximation} \\ \dfrac{(1+\lambda \frac{P}{U})^2}{(z_\epsilon + \frac{\gamma(z_\epsilon^2-1)}{6})^2 k \frac{P}{U}}, & \text{Normal power transformation.} \end{cases}$$

On Figure 1 are shown graphics of different values of λ, [4]. The function reaches minimum, which can be calculated solving the differential equation

$$\frac{\partial M}{\partial P} = \frac{1}{z_\epsilon^2 k} U(-\frac{U}{P^2} + \frac{\lambda^2}{U}) = 0$$

The last equality is fulfilled when $P = \dfrac{U}{\lambda}$, and consequently,

$$M_{min} = 4\frac{1}{z_\epsilon^2 k}\lambda U.$$

7. Conclusions

The choice of the distribution of the size of the individual claims is of importance for the results of the considered methods. When there is high asymmetry, the size of the errors from the approximation is rising up significantly. According to the considered cases, the most stable method is the normal power transformation. For the positive values of the coefficient of asymmetry of the distribution of the total value of the claims, the normal approximation gives lower value for the estimation of the reserves than the Normal power transformation. The modeling of the claim process gives significantly influence of the results from the risk estimation.

REFERENCES

[1] A. Oliveri, E. Pitacco. Introduction to Insurance Mathematics, Springer, 2011.

[2] D. G. Hart, R. A. Buchanan, B. A. Howe Actuarial Practice in General Insurance. Institute of Actuaries of Australia, Sydney, 1996

[3] E. Raeva, V. Pavlov. Some Approaches for Modeling Claims Process in General Insurance, 2015.

[4] E. Raeva, V. Pavlov. On The Risk Modeling In General Insirance, 2015.

[5] I. B. Hossack, J. H. Pollard, B. Zenhwirth. Introductory statistics with applications in general insurance. Cambridge, Cambridge University Press, 1983.

[6] J. H. Pollard. A handbook of numerical and statistical techniques. Cambridge, Cambridge University Press, 1977.

Elitsa Raeva
e-mail: `eraeva@uni-ruse.bg`
Velizar Pavlov
e-mail: `vpavlov@uni-ruse.bg`
Depratment of Applied Mathematics and Statistics
University of Ruse
Studentska Str 8, 7017 Ruse, Bulgaria