ISBN: 978-99955-45-42-0
УДК: 005.5:368

Датум пријема рада: 17.05.2023.
Датум прихватања рада: 13.06.2023.
Оригинални научни рад

ЗНАЧАЈ ПОУЗДАНОСТИ СТАТИСТИЧКИХ РАСПОДЕЛА ЗА ФОРМИРАЊЕ ПРЕМИЈЕ У НЕЖИВОТНОМ ОСИГУРАЊУ

THE SIGNIFICANCE OF RELIABILITY OF STATISTICAL DISTRIBUTIONS FOR PREMIUM CALCULATION IN NON-LIFE INSURANCE

Јелена Станојевић
Универзитет у Београду, Економски факултет, Београд, Србија
jelena.stanojevic@ekof.bg.ac.rs
ORCID: 0000-0001-5668-5297

Весна Рајић
Универзитет у Београду, Економски факултет, Београд, Србија
vesna.rajic@ekof.bg.ac.rs
ORCID: 0000-0002-4566-0147

Марија Копривица
Универзитет у Београду, Економски факултет, Београд, Србија
marija.koprivica@ekof.bg.ac.rs
ORCID: 0000-0003-4239-2252

Апстракт: *Предмет рада је анализа најзначајнијих расподела које се примењују у актуарству и осигурању, као и њихових карактеристика у виду очекиване вредности, варијансе, асиметрије и спљоштености. Циљ рада је да укаже на значај поузданости статистичких расподела у израчунавању премије у неживотном осигурању. Основни резултат рада је одређивање најбоље расподеле агрегатног износа штета која описује дати скуп података који је разматран и потврђује важност правилног избора расподеле у израчунавању премије, што је била основна претпоставка рада.*

Кључне ријечи: *расподеле броја и износа штета, оцена расподеле, израчунавање премије, неживотно осигурање*

JEL класификација: *C46, C18, G22*

Abstract: *The subject of the paper is analysis of the most important distributions which are applied in actuarial science and insurance, and also their characteristics in*

the form of expected value, variance, skewness and excess kurtosis. The aim of the paper is to emphasize the importance of reliability of statistical distributions in calculating premiums in non-life insurance. The main result in the paper is determination of the best distribution of the aggregate claim amount that describes the given dataset and confirms the importance of choosing the correct distribution in premium calculation, which was the main assumption of the paper.

Key Words: *distributions of number and amount of claims, distribution fitting, premium calculation, non-life insurance*

JEL classification: *C46, C18, G22*

1. INTRODUCTION

The goal of insurance is to spread risks through a community of vulnerable individuals. It plays a significant role in promoting financial stability by mitigating uncertainty, facilitating exchange and trade, etc. Non-life insurance covers risks related to property and liability, providing financial compensation in the event of damage or loss. In order to fulfill the insurer's obligations towards insurance beneficiaries, it is essential to accumulate adequate insurance premiums. The insurance premium is affected by the risk, the sum insured, the duration of the insurance and the interest rate (Kočović, Rakonjac-Antić, Koprivica, Šulejić, 2021). Non-life insurance contracts are most often concluded for one year, which affects the dynamics and maturity of insurance premium placement.

In this paper we examine the importance of determining the appropriate probability distribution in claim data analysis, which is confirmed through an example of premium calculation in non-life insurance. The data on claims in motor vehicle insurance (casco) have previously been used in Jovović (2015). We covered 6976 insurance policies in 2013. The paper is organised into four parts. The first part is introduction. In the second part of the paper we consider in detail the *ad hoc* method of premium principles, as one of the three methods for determining the price in non-life insurance. We also emphasize importance of applying different statistical distributions in insurance and consider decision making which distribution fits dataset the best. The third part of the paper contains an overview of the most important and frequently used distributions in a non-life insurance with their characteristics which we use for premium calculation. In the fourth part of the paper we present results of data analysis, which determined the best-fitting distribution for the considered dataset and confirmed the significance of using statistical distributions for insurance premium calculation.

2. THE SIGNIFICANCE OF RELIABILTY OF STATISTICAL DISTRIBUTION

One important aspect of the significance of the corect determination of statistical distribution in the field of insurance relates to premium calculation in non-life insurance. Insurance premium represents the price of insurance. Namely, loss values are real random variables ($X \in [0, +\infty)$), and it is of great importance for insurers to

determine the distribution functions of the size and frequency of losses when calculating premiums. One of the main topics in actuarial science is the definition and selection of real principles of insurance premium calculation. In this paper we consider calculation of the technical premium, which is calculated based on predictions of future insurable events. The literature describes three methods for premium calculation: the ad hoc method (a potential premium principle is defined based on the loss distribution, and then evaluated if it possesses appropriate properties), the characterization method (the desirable properties of the premium pinciple are first defined and then all premium principles that satisfy them are identified) and the economic method (first, it is important to know economic theory and then to define principles of premium calculation) (Young, 2004, p. 1).

2.1. Ad hoc method in determining insurance premium

Let (Ω, F, p) be the probability space and χ denote the set of nonnegative random variables defined on that space, that is collection of insurance-loss random variables, usually called insurance risks. Let X denote a member of χ. Finally, let P denote the premium principle. Further, we list well known premium principles under ad hoc method of insurance premium calculation (Young, 2004, p. 3-5).

1. *Net Premium Principle*: $P[X] = EX$.

This is the widely applied premium principle in the literature because actuaries often assume if the insurer sells a sufficient number of identically distributed and independent policies, then the risk effectively disappears (Bowers, Gerber, Hickman, Jones and Nesbitt, 1997).

2. *Expected Value Premium Principle*: $P[X] = (1 + \theta)\, EX,\ \theta > 0$.

This premium principle includes a risk load which is proportional to the expected value of the risk.

3. *Standard Deviation Premium Principle*: $P[X] = EX + \alpha\sqrt{VarX},\ \alpha > 0$.

Like previous two principles, this premium principle is also based on the Net Premium Principle and includes a risk load which is proportional to the standard deviation of the risk.

4. *Variance Premium Principle*: $P[X] = EX + \beta\, VarX,\ \beta > 0$.

This premium principle includes a risk load which is proportional to the variance of the risk (Bühlmann, 2007).

5. *Exponential Premium Principle*: $P[X] = \left(\frac{1}{\alpha}\right) \ln E[e^{\alpha X}],\ \alpha > 0$.

This principle satisfies many nice properties, for example additivity with respect to independent risks.

6. *Esscher Premium Principle*: $P[X] = (\frac{E[X\, e^{Z}]}{E[e^{Z}]})$, for some random variable Z.

In the literature this pinciple was considered in the case $Z = f(X)$ for some arbitrary function f (Heilmann, 1989) and the case $e^Z = 1 - e^{-\lambda X}$, $\lambda > 0$ (Kamps, 1998). Some authors define this principle with $Z = h\,X$, $h > 0$.

7. *Wang's Premium Principle*: $P[X] = \int_0^\infty g[S_X(t)]dt$, where g is an increasing, concave function that maps $[0,1]$ onto $[0,1]$.

The function g is called a distortion and $g[S_X(t)]$ is called a distorted (tail) probability.

8. *Proportional Hazards Premium Principle*: $P[X] = \int_0^\infty [S_X(t)]^c dt$, for some $0 < c < 1$.

This is a special case of Wang's Premium Principle, for $g(x) = x^c$, $0 < c < 1$.

9. *Principle of Equivalent Utility*: $P[X]$ is a solution of equation $u(\omega) = E[u(\omega - X + P)]$, where ω is the initial wealth of the insurer, u in an increasing, concave utility function of wealth of the insurer, P is a minimum premium that the insurer is willing to accept in exchange for insuring the risk X.

Special case of this principle is Exponential Premium Principle, for utility function: $u(\omega) = -e^{-\alpha\omega}$, $\alpha > 0$. Also, if the premium P is the solution of the equation $E[u(P - X)] = u(0)$, then we use terminology *principle of zero-utility* (Bühlmann, 2007).

10. *Swiss Premium Principle*: the premium P is a solution of the equation $E[u(X - pP)] = u((1 - p)P)$, for some $p \in [0,1]$ and some increasing, convex function u.

This principle is a generalization of the principle of zero-utility, when we change function $u(x)$ with the function $-u(-x)$) and for $p = 1$.

11. *Dutch Premium Principle*: $P[X] = EX + \theta E[(X - \alpha\, EX)_+]$, $\alpha \geq 1$, $0 < \theta \leq 1$.

This principle was introduced in the literature (Van Heerwaarden and Kaas, 1992). Also this principle is extended to reinsurance (Hürlimann, 1994) .

2.2. Determing distribution which fits the dataset the best

The main question is how to find the best theoretical distribution for given dataset, because knowledge about the underlying data distribution is the main step for data modelling and has many applications. The histogram is the easiest way for obtaining a visual representation of the underlying distribution of random variables. Various plots can be created such as probability distribution function or cumulative distribution function (pdf/cdf), and the Quantile-Quantile plot (QQ plot), for the candidate distribution. Because pdf is a fundamental concept in statistics, it is important to specify the function f that gives a natural description of the distribution of a given random variable X. Exploring fundamental characteristics of the data, such as skewness, kurtosis, outliers, distribution shape, univariate, bimodal, etc. is also

important in data modelling, because by understanding that characteristics it becomes easier to decide which models are better suited for the data.

The four steps to determine the theoretical distributions are: computing density and weights from a histogram, estimating distribution parameters from the data, checking goodness-of-fit, and s**electing the best theoretical distribution.**

3. THE MOST COMMONLY USED DISTRIBUTIONS IN NON-LIFE INSURANCE AND THEIR CHARACTERISTICS

The probabilistic model representing the total (aggregate) amount of claims combines two components: the distribution of the number of claims and the distribution of the amount of individual claims. In this section, we will present the most important theoretical distribution models commonly used in non-life insurance, as identified in (Jovović, 2015). Also, we will give the main characteristics of that distributions, including parameters, important measures of the distibutions: mean (μ) and variance (σ^2) values, skewness (γ_1) and excess kurtosis (γ_2), and coefficient of variation (C_V), which shows the variability in percentages of the expected value.

3.1. Modeling the number of claims

The first step for modeling the aggregate amount of claims is the selection of the appropriate theoretical distribution of the number of claims. That is a discrete, non-negative random variable N. The most commonly used distributions in that purpose are Bernoulli, binomial, Poisson and negative binomial distributions.

The simplest type of distribution that can be used for modeling number of claims is *Bernoulli distribution*, $Ber(p)$. The appropriate probability mass function is in the form:

$$P(N = k) = p^k(1-p)^{1-k}, k \in \{0,1\},$$

for $0 < p < 1$, where p represents the probability of a claim event realization and it is evaluated based on the average frequency of claims. Appropriate measures of this distribution are:

$$\mu = E(N) = p,\ \sigma^2 = Var(N) = pq,\ q = 1-p$$

$$\gamma_1 = E(\frac{N-\mu}{\sigma})^3 = \frac{\mu_3}{\sigma^3} = \frac{q-p}{\sqrt{pq}}, \gamma_2 = E(\frac{N-\mu}{\sigma})^4 - 3 = \frac{\mu_4}{\sigma^4} - 3 = \frac{1-6pq}{pq}, C_V = \frac{\sigma}{\mu} = \sqrt{\frac{q}{p}}$$

where we use the notation $\mu_k = E(N-\mu)^k$. We cannot make any general statements about the skewness and kurtosis of the Bernoulli distribution because both of these measures depend on the value of the parameter p.

In non-life insurance, the Bernoulli distribution is suitable for modeling the number of claims where only one insured event can arise per policy during the insurance coverage period, what is a highly restrictive condition.

If the maximum number of claims per policy is known and it is n, each of them has the same probability of occurrence $0 < p < 1$ and are independent, then the random variable N follows a *binomial distribution*, $Bin(n, p)$. The appropriate probability mass function is in the form:

$$P(N = k) = \binom{n}{k} p^k (1-p)^{n-k}, k = 0,1, \dots, n.$$

The selection of the parameter n is based on the conditions of the specific insurance contract, while the parameter p is evaluated based on the average frequency of claims. This distribution is suitable for modeling random variables for which it is not possible or it is not logical, to assume that, with a positive probability, they take a value above a certain level (such as for example, the number of damages on one car during the year). That is obvious because of the limited domain of definition for this distribution. Appropriate measures of this distribution are:

$$\mu = E(N) = np, \sigma^2 = Var(N) = npq, \ q = 1 - p$$

$$\gamma_1 = E(\frac{N-\mu}{\sigma})^3 = \frac{\mu_3}{\sigma^3} = \frac{q-p}{\sqrt{npq}}, \gamma_2 = E(\frac{N-\mu}{\sigma})^4 - 3 = \frac{\mu_4}{\sigma^4} - 3 = \frac{1-6pq}{npq}, C_V = \frac{\sigma}{\mu} = \sqrt{\frac{q}{np}}$$

This distribution corresponds to situations in which the sample variance is smaller than the sample mean.

In the case when n is large enough and p is sufficiently small, binomial distribution can be approximated with *Poisson distribution*, $Poi(\lambda)$ with parameter $\lambda = np$. The appropriate probability mass function is in the form:

$$P(N = k) = e^{-\lambda} \frac{\lambda^k}{k!}, k = 0,1,2, \dots$$

This distribution, compared to the binomial distribution, usually fits data in non-life insurance more accurately. The important advantages are that Poisson distribution requires the estimation of only one parameter and the fact that its application does not require prior determination of the maximum number of damages per insurance policy. Appropriate measures of this distribution are:

$$\mu = E(N) = \lambda, \sigma^2 = Var(N) = \lambda$$

$$\gamma_1 = E(\frac{N-\mu}{\sigma})^3 = \frac{\mu_3}{\sigma^3} = \frac{1}{\sqrt{\lambda}}, \gamma_2 = E(\frac{N-\mu}{\sigma})^4 - 3 = \frac{\mu_4}{\sigma^4} - 3 = \frac{1}{\lambda}, C_V = \frac{\sigma}{\mu} = \frac{1}{\sqrt{\lambda}}$$

From the formulas of expected value and variance for Poisson distribution we may conclude that it is not adequate distribution in a situation where the variance of a

given variable exceeds its expected value. The parameter λ, as the expected number of claims per insurance policy, can be estimated on the base of the average frequency of claims. However, the implicit assumption is that all policies of the observed portfolio are homogeneous in terms of frequency of claims.

The frequency of claims for each of the policies may be lower or higher than the calculated average value. Therefore, a more reliable approach is that λ is a random variable. If λ follows, for example, a gamma distribution with parameters r and $\frac{p}{1-p}$, the variable N, number of claims, follows a *negative binomial distribution*, $NB(r,p)$. The appropriate probability mass function is in the form:

$$P(N=k)=\binom{k+r-1}{k}p^r(1-p)^k, k=0,1,2,\dots$$

where $r>0$ and $0<p<1$ are parameters of the distribution. Appropriate measures of this distribution are:

$$\mu=E(N)=\frac{r(1-p)}{p},\ \sigma^2=Var(N)=\frac{r(1-p)}{p^2}$$

$$\gamma_1=E(\frac{N-\mu}{\sigma})^3=\frac{\mu_3}{\sigma^3}=\frac{1+p}{\sqrt{pr}},\gamma_2=E(\frac{N-\mu}{\sigma})^4-3=\frac{\mu_4}{\sigma^4}-3$$

$$=\frac{6}{r}+\frac{p^2}{r(1-p)},C_V=\frac{\sigma}{\mu}=\frac{1}{\sqrt{r(1-p)}}$$

As negative binomial distribution allows the relation $Var(N)>E(N)$, it is relevant distribution in the case when the variance of the number of claims is greater than its mean. Also it is applied when the portfolio includes policies with a higher number of claims per year, because the tail of the negative binomial distribution for $r>1$ decreases more slowly than the tail of the geometric distribution, which is a special case of negative binomial distribution for $r=1$ and represents the natural boundary between the distributions of heavy-tailed and light-tailed discrete random variables (Jovović, 2015).

3.2. Modeling the amount of individual claims

Absolutely continuous probability distributions are used in the collective risk model for the purposes of modeling the variable amount (intensity) of losses X. The usual domain of definition of the used distributions is $(0,\infty]$, although the set of possible values of a given variable is limited in reality and there are assumptions that the expected value of the amount of losses is finite, while the dispersion can be infinite. Distributions that desribe data on the amount of claims in non-life insurance, which are the most commonly used in practice are gamma, exponential, log-normal, Weibull, Pareto, Burr, log-gamma and inverse Gaussian distributions (Jovović , 2015).

The random variable X follows *gamma distribution*, $Gamma(\alpha,\beta)$ if the pdf is in the form:

$$f(x) = \begin{cases} \frac{1}{\beta^{\alpha}\Gamma(\alpha)} x^{\alpha-1} e^{-\frac{x}{\beta}}, & x \geq 0 \\ 0, & x < 0 \end{cases}$$

where $\Gamma(\alpha) = \int_0^{\infty} x^{\alpha-1} e^{-x} dx$ is a gamma function and $\alpha > 0$ is a shape parameter and $\beta > 0$ is a scale parameter of the distribution. This distribution has nice properties, which can be a reason for its application in modeling losses. Although rarely used to describe losses by itself, the gamma distribution is often used in combination with other distributions for that purpose. Appropriate measures of the gamma distribution are:

$$\mu = E(X) = \alpha\beta, \sigma^2 = Var(X) = \alpha\beta^2$$

$$\gamma_1 = E(\frac{X-\mu}{\sigma})^3 = \frac{\mu_3}{\sigma^3} = \frac{2}{\sqrt{\alpha}}, \gamma_2 = E(\frac{X-\mu}{\sigma})^4 - 3 = \frac{\mu_4}{\sigma^4} - 3 = \frac{6}{\alpha}, C_V = \frac{\sigma}{\mu} = \frac{1}{\sqrt{\alpha}}$$

Special case of the gamma distribution for $\alpha = 1$ is *exponential distribution*, $Exp(\lambda)$, with the pdf:

$$f(x) = \begin{cases} \lambda e^{-\lambda x}, & x \geq 0 \\ 0, & x < 0 \end{cases}$$

This distribution is a typical representative of light-tailed distributions because its right tail declines at an exponential rate, $\bar{F}(x) = P(X > x) = e^{-\lambda x}$, $x \geq 0$. Such a distribution is suitable for modeling the variable amount of losses, due to the existence of its moment generating function in the neighborhood of zero, from theoretical point of view.

These distibutions have relatively wide application in the motor insurance domain. However, in non-life insurance, large and catastrophic losses are possible, which correspond to distributions with a heavy tail, for which the moment generating function is not defined in the neighborhood of zero.

With exponential transformation of variables that have one of the usual light-tailed distributions it is possible to derive certain classes of heavy-tailed distributions. Such is the case with the *log-normal distribution*, $LN(m, \sigma^2)$ of the random variable $X = e^Y$, where Y has a normal distribution $N(m, \sigma^2)$. The appropriate pdf is in the form:

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

where $m \in R$ is a shape parameter and $\sigma > 0$ is a scale parameter of the distribution. By logarithmizing the data it is posible to transform log-normal distribution into a normal distribution, which can be more convenient for data analysis in practice. Appropriate measures of the log-normal distribution are:

$$\mu = E(X) = e^{m+\frac{\sigma^2}{2}}, \sigma^2 = Var(X) = e^{2m+\sigma^2}(e^{\sigma^2} - 1)$$

$$\gamma_1 = E(\frac{X-\mu}{\sigma})^3 = \frac{\mu_3}{\sigma^3} = \sqrt{e^{\sigma^2}-1}\ \left(e^{\sigma^2}+2\right), \gamma_2 = E(\frac{X-\mu}{\sigma})^4 - 3 = \frac{\mu_4}{\sigma^4} - 3$$
$$= e^{4\sigma^2} + 2e^{3\sigma^2} + 3e^{2\sigma^2} - 6$$

$$C_V = \frac{\sigma}{\mu} = \sqrt{e^{\sigma^2}-1}$$

Generalisation of exponential distribution is the *Weibull distribution*, $Weib(k,\lambda)$, with the pdf:

$$f(x) = \begin{cases} \frac{k}{\lambda}(\frac{x}{\lambda})^{k-1} e^{-(x/\lambda)^k}, & x \geq 0 \\ 0, & x < 0 \end{cases}$$

where $k > 0$ is a shape parameter and $\lambda > 0$ is a scale parameter of the distribution. In a case when $k = 1$ the Weibull distribution reduces to the exponential distribution. The tail of the Weibull distribution is lighter than the exponential tail for $k > 1$, or heavier for $k < 1$. Appropriate measures of the Weibull distribution are:

$$\mu = E(X) = \lambda\,\Gamma(1+\tfrac{1}{k}),\ \sigma^2 = Var(X) = \lambda^2\,(\Gamma\left(1+\tfrac{2}{k}\right) - \Gamma^2\left(1+\tfrac{1}{k}\right))$$

$$\gamma_1 = E(\frac{X-\mu}{\sigma})^3 = \frac{\mu_3}{\sigma^3} = \frac{\Gamma\left(1+\frac{3}{k}\right)\lambda^3 - 3\mu\sigma^2 - \mu^3}{\sigma^3},$$

$$\gamma_2 = E(\frac{X-\mu}{\sigma})^4 - 3 = \frac{\mu_4}{\sigma^4} - 3 = \frac{\Gamma\left(1+\frac{4}{k}\right)\lambda^4 - 4\,\gamma_1\sigma^3\,\mu - 6\sigma^2\mu^2 - \mu^4}{\sigma^4} - 3$$

$$C_V = \frac{\sigma}{\mu} = \frac{\sqrt{\Gamma\left(1+\frac{2}{k}\right) - \Gamma^2\left(1+\frac{1}{k}\right)}}{\Gamma(1+\frac{1}{k})}$$

One of the most important distributions that find its application in non-life insurance is the *Pareto (Type I) distribution*, $Par(\alpha,\beta)$ with the pdf:

$$f(x) = \begin{cases} \frac{\alpha\beta^\alpha}{x^{\alpha+1}}, & x \geq \beta \\ 0, & x < \beta \end{cases}$$

where $\alpha > 0$ is a shape parameter and $\beta > 0$ is a scale parameter of the distribution. The Pareto distribution is used primarily for modeling extremely large losses. For this distribution it is valid: $E(X^k) = \infty$ for $k > \alpha$. That means that for small value α, variance and even more mean value are not finite, which is a common situation in reinsurance. The Pareto distribution, like the exponential one, is not adequate in many practical situations because of the monotonically decreasing tail (Burnecki, Misiorek and Weron, 2005, p. 300). Appropriate measures of the Pareto distribution are:

$$\mu = E(X) = \begin{cases} \infty, & \alpha \leq 1 \\ \frac{\alpha\beta}{\alpha-1}, & \alpha > 1 \end{cases},\ \sigma^2 = Var(X) = \begin{cases} \infty, & \alpha \leq 2 \\ \frac{\alpha\beta^2}{(\alpha-2)(\alpha-1)^2}, & \alpha > 2 \end{cases}$$

$$\gamma_1 = E(\frac{X-\mu}{\sigma})^3 = \frac{\mu_3}{\sigma^3} = \frac{2(\alpha+1)}{\alpha-3}\sqrt{\frac{\alpha-2}{\alpha}}, \alpha > 3, \gamma_2 = E(\frac{X-\mu}{\sigma})^4 - 3$$
$$= \frac{6(\alpha^3+\alpha^2-6\alpha-2)}{\alpha(\alpha-3)(\alpha-4)}, \alpha > 4$$

$$C_V = \frac{\sigma}{\mu} = \frac{1}{\sqrt{\alpha(\alpha-2)}}, \alpha > 2$$

Widely applied distribution in non-life insurance is the *Burr (Type XII) distribution*, $Burr(\alpha, \beta)$ with the pdf:

$$f(x) = \frac{\alpha\beta x^{\alpha-1}}{(1+x^{\alpha})^{\beta+1}}, x > 0$$

where $\alpha > 0$ and $\beta > 0$ are shape parameters of the distribution. With Burr distribution it is possible to overcome problems that arise during application Pareto distribution. Appropriate measures of the Burr distribution are:

$$\mu = E(X) = \beta\, \mathrm{B}\left(\beta - \frac{1}{\alpha}, 1 + \frac{1}{\alpha}\right), \sigma^2 = Var(X) = -\mu_1^2 + \mu_2$$

$$\gamma_1 = E(\frac{X-\mu}{\sigma})^3 = \frac{\mu_3}{\sigma^3} = \frac{2\mu_1^3 - 3\mu_1\mu_2 + \mu_3}{(-\mu_1^2+\mu_2)^{3/2}}, \gamma_2 = E(\frac{X-\mu}{\sigma})^4 - 3 = \frac{\mu_4}{\sigma^4} - 3$$
$$= \frac{-3\mu_1^4 + 6\mu_1^2\mu_2 - 4\mu_1\mu_3 + \mu_4}{(-\mu_1^2+\mu_2)^2} - 3$$

$$C_V = \frac{\sigma}{\mu} = \frac{\sqrt{-\mu_1^2+\mu_2}}{\beta\, \mathrm{B}(\beta - \frac{1}{\alpha}, 1 + \frac{1}{\alpha})}$$

where $\mu_r = \beta\, \mathrm{B}(\beta - \frac{r}{\alpha}, 1 + \frac{r}{\alpha})$ and $B(\,,\,)$ is a beta function. It's useful to know the relation between Pareto and Burr distribution. For that purpose we will give new parametrisation and tails of the Pareto and Burr distributions. Namely, if we consider a random variable Y which follows the Pareto distribution with parameters $\alpha > 0$ and $\theta > 0$ and tail which is in the form

$$\bar{F}(x) = \begin{cases} (\frac{\theta}{\theta+x})^{\alpha}, & x > 0 \\ 0, & x \le 0 \end{cases}$$

with the introduction of a new parameter $\tau > 0$ and new variable $X = Y^{1/\tau}$, then X follows Burr distribution with the tail

$$\bar{F}(x) = \begin{cases} (\frac{\theta}{\theta+x^{\tau}})^{\alpha}, & x > 0 \\ 0, & x \le 0 \end{cases}$$

The distribution of the random variable $X = e^Y + (\theta - 1)$, where Y follows gamma distribution $Gamma(\alpha, \beta)$, is a new *log-gamma distribution*, $LogGamma(\alpha, \beta, \theta)$ with the pdf:

$$f(x) = \frac{1}{\beta^{\alpha}\Gamma(\alpha)} (x - \theta + 1)^{-(1+\frac{1}{\beta})} (\ln(x - \theta + 1))^{\alpha-1}, x \geq \theta$$

where $\alpha, \beta, \theta > 0$ are parameters of the distribution. This distribution is useful when dealing with data that has extreme values, where an exponential or logarithmic transformation is required to normalize the data (Consul and Jain, 1971, p. 100). Appropriate measures of the distribution are:

$$\mu = E(X) = (1-\beta)^{-\alpha} + \theta - 1, \beta < 1, \sigma^2 = Var(X) = (1-2\beta)^{-\alpha} - (1-\beta)^{-2\alpha}, \beta < \frac{1}{2}$$

$$\gamma_1 = E(\frac{X-\mu}{\sigma})^3 = \frac{\mu_3}{\sigma^3} = \frac{(1-3\beta)^{-\alpha} - 3(1-3\beta+2\beta^2)^{-\alpha} + 2(1-\beta)^{-3\alpha}}{\sigma^3}, \beta < \frac{1}{3}$$

$$\gamma_2 = E\left(\frac{X-\mu}{\sigma}\right)^4 - 3 = \frac{\mu_4}{\sigma^4} - 3 = \frac{(1-4\beta)^{-\alpha} - 4(1-4\beta+3\beta^2)^{-\alpha} + 6(1-2\beta)^{-\alpha}(1-\beta)^{-2\alpha} - 3(1-\beta)^{4\alpha}}{\sigma^4}, \beta < \frac{1}{4}$$

$$C_V = \frac{\sigma}{\mu} = \frac{\sqrt{(1-2\beta)^{-\alpha} - (1-\beta)^{-2\alpha}}}{(1-\beta)^{-\alpha} + \theta - 1}, \beta < \frac{1}{2}$$

Application of log-gamma distribution in practice can be difficult due to the absence of variance and expected values depending on the parameters.

The *Inverse Gaussian distribution*, $InvGauss(\alpha, \beta)$ with two parameters has the pdf in the form:

$$f(x) = \frac{\alpha}{\sqrt{2\pi x^3}} e^{-\frac{(\alpha-\beta x)^2}{2x}}, x > 0$$

where $\alpha > 0$ is a location parameter and $\beta > 0$ is a scale parameter of the distribution. Appropriate measures of the distribution are:

$$\mu = E(X) = \frac{\alpha}{\beta}, \sigma^2 = Var(X) = \frac{\alpha}{\beta^3},$$

$$\gamma_1 = E(\frac{X-\mu}{\sigma})^3 = \frac{\mu_3}{\sigma^3} = \frac{3}{\sqrt{\alpha\beta}}, \gamma_2 = E(\frac{X-\mu}{\sigma})^4 - 3 = \frac{\mu_4}{\sigma^4} - 3 = \frac{15}{\alpha\beta}, C_V = \frac{\sigma}{\mu} = \frac{1}{\sqrt{\alpha\beta}}$$

This distribution can be useful in practice due to its favorable properties. Specifically, if X follows the inverse Gaussian distribution, then $Y = \sqrt{\alpha\beta}\left(\frac{\beta X}{\alpha} - 1\right)$ and $Z = \frac{1}{2}\frac{1}{\sqrt{\alpha\beta}} + \sqrt{\alpha\beta}\ln\frac{\beta X}{\alpha}$ follow an approximately standard normal distribution, $N(0,1)$.

At the end of this section we need to summarize the final step in statistical analysis for determining the expected value used to calculate insurance premium. This involves determining the compound distribution of the aggregate amount of all claims for a given portfolio, during the observed time period by combining distributions of the number and amount of individual claims. If N denotes the number of claims in the considered period and X_i denotes the amount of individual claim, then S is an aggregate amount of all claims and can be expressed as the sum $S = \sum_{i=1}^{N} X_i$, where $S = 0$ for $N = 0$.

The collective risk model assumes that the amounts of claims $X_1, X_2, \ldots$ are i.i.d. random variables, and that they are also independent of the number of claims N. Based on the complete probability formula, the next relation is satisfied

$$F_S(x) = P(S \le x) = \sum_{n=0}^{\infty} P(S \le x | N = n)\, P(N = n)$$

Let F_X^{*n} denotes the n-th convolution of distribution function of individual claim amount $F_X(x) = P(X \le x)$,

$$F_X^{*n}(x) = P(S \le x | N = n) = P(\sum_{i=1}^{n} X_i \le x)$$

where $F^{*0}(x) = 1$ for $x \ge 0$ and $F^{*0}(x) = 0$ for $x < 0$, by convention. Then, we have $F_S(x) = \sum_{n=0}^{\infty} F_X^{*n}(x)\, P(N = n)$ for $x \ge 0$. On the base of conditional expectation and conditional variance, under which for the two random variables Y and Z, it is satisfied: $E(Y) = E(E(Y|Z))$ and $Var(Y) = E\big(Var(Y|Z)\big) + Var(E(Y|Z))$, we may calculate expectation and variance of aggregate amount of claims S

$$E(S) = E(X)\, E(N)\,,\ Var(S) = E(N)Var(X) + \big(E(X)\big)^2 Var(N)$$

where $Var(X) = E(X^2) - (E(X))^2$.

4. DATA ANALYSIS

In this section we will calculate premium values based on three most commonly mentioned premium principles in the literature, and compare them.

We observe data on claims under the motor vehicle insurance previously employed in Jovović (2015). We covered 6976 insurance policies in 2013. Using EasyFit it can be shown that the number of claims has a Poisson distribution with parameter $\lambda = 0.08686$. This means that for 100,000 policies, an average of 8,686 claims can be expected during the year.

We found that the amounts of losses can be well described by the log-gamma distribution with estimated parameter values of $\alpha = 105.50$, $\beta = 0.107285$ and $\theta = 1$. The value of the Kolmogorov-Smirnov test statistic is 0.05480, and the corresponding p value is 0.05056, which suggests that, at a significance level of 0.05, we cannot reject the null hypothesis that the data follows the log-gamma distribution with the estimated parameters.

Further, based on the formulas in Section 3, we calculated the mean and variance of the amount of loss X, resulting in $\mu = E(X) = 158{,}414.9661$ and $Var(X) = 91{,}360{,}275{,}013$; then the mean and variance of aggregate amount of claims S are: $E(S) = E(X)\,E(N) = 13{,}759.92396$, $Var(S) = E(N)Var(X) + \big(E(X)\big)^2 Var(N) = 10{,}115{,}331{,}375$.

Now we calculate premium values based on the first three premium principles. Results are shown in Table 1. In all formulas we assume that θ, or α, takes one of the three values: 0.1, 0.8 and 1. We will call that coefficient, the risk aversion coefficient. The first value of the coefficient θ (also for α) is chosen according to the work (Lesmana, Wulandari, Napitupulu and Supian, 2018), second coefficient was chosen in order to observe the effect of increasing the risk aversion coefficient on the amount of the premium, and the third coefficient was chosen in order to see what the premium would be at the maximum value of this coefficient.

Table 1. Calculated insurance premium (in RSD)

Premium Principle	Risk aversion coefficient 0.1	Risk aversion coefficient 0.8	Risk aversion coefficient 1
Net Premium Principle	13,759.92396	13,759.92396	13,759.92396
Expected Value Principle	15,135.91635	24,767.86312	27,519.84791
Standard Deviation Principle	23,817.42433	94,219.92694	114,334.9277

Source: *Authors calculations*

Based on Table 1 the Net Premium Principle is not adequate as it does not incorporate a risk load to account for potential deviations of actual from expected losses. If we look at the Expected Value Principle, increasing the risk aversion coefficient from 0.1 to 0.8 leads to an increase in the premium 1.64 times. At the maximum value of this coefficient, the premium increases 1.82 times. With the Standard Deviation Principle, these ratios are much higher. Increasing risk aversion from 0.1 to 0.8 leads to an increase in the premium 3.96 times. At the maximum value of this coefficient, the premium increases 4.8 times. By increasing these coefficients, the risk load also increases. The conducted analysis shows that the amount of insurance premium can significantly vary depending on the selected principles and coefficients. The actuary is responsible for determining insurance premiums that are sufficient to cover potential losses and maintain the insurance company's solvency. Therefore, the experience and

expertise of actuaries are crucial for determining the amount of risk load necessary for the insurance premium to be adequate.

CONCLUSION

In this paper we pointed out the importance of determing the best fitting distribution for considerid dataset for a further manipulation of that data, in our case for non-life insurance premium calculation. Also, we presented a methodological framework to identify an appropriate probability model that best describes the frequency and amount of insurance claims and applied it to analyze the data of motor vehicle (casco) insurance. In empirical analysis, the number of claims can be modeled with a Poisson distribution, while the amount of claims can be modeled with a log-gamma distribution. At the end we gave a comparison of premiums calculated according to four premium principles. We concluded that different principles lead to different insurance premium amounts, and that the actuary should decide which premium principle is the best in a concrete case.

REFERENCES

[1] Bowers, N. L., Gerber, H. U., Hickman, J. C., Jones, D. A., Nesbitt, C. J. (1997). *Actuarial Mathematics*, 2nd edition, Society of Actuaries, Schaumburg, IL.

[2] Bühlmann, H. (2007). *Mathematical methods in risk theory* (Vol. 172). Springer Science & Business Media.

[3] Burnecki, K., Misiorek, A., Weron, R. (2005). 13 Loss Distributions. *Statistical tools for finance and insurance*, Berlin: Springer-Verlang, Ch. 13, 289-318.

[4] Consul, P.C., Jain, G.C. (1971). On the log-gamma distribution and its properties. *Statistische Hefte*, 12(2), 100-106.

[5] Heilmann, W. R. (1989). Decision theoretic foundations of credibility theory. *Insurance: Mathematics and Economics*, *8*(1), 77-95.

[6] Hürlimann, W. (1994). A note on experience rating, reinsurance and premium principles. *Insurance: Mathematics and Economics*, *14*(3), 197-204.

[7] Jovović, M. (2015). *Merenje rizika pri utvrđivanju solventnosti neživotnih osiguravača*. Doctoral dissertation, Belgrade: Faculty of Economics, University of Belgrade.

[8] Kamps, U. (1998). On a class of premium principles including the Esscher principle. *Scandinavian Actuarial Journal*, *1998*(1), 75-80.

[9] Kočović, J., Rakonjac-Antić T., Koprivica M., Šulejić P. (2021). *Osiguranje u teoriji i praksi*. Belgrade: Faculty of Economics, University of Belgrade.

[10] Lesmana, E., Wulandari, R., Napitupulu, H. and Supian, S. (2018). Model estimation of claim risk and premium for motor vehicle insurance by using Bayesian method. In *IOP Conference Series: Materials Science and Engineering* (Vol. 300, No. 1, p. 012027). IOP Publishing.

[11] Young, V. R. (2004). Premium Principles. *Reprodiced from Encyclopedia of Acturial Science*, 1-9.

[12] Van Heerwaarden, A. E., Kaas, R. (1992). The Dutch premium principle. *Insurance: Mathematics and Economics*, *11*(2), 129-133.

[13] Wang, S. (1995). Insurance pricing and increased limits ratemaking by proportional hazards transforms. *Insurance: Mathematics and Economics*, *17*(1), 43-54.

SUMMARY

This paper deals with the analysis of the major distributions used in actuarial science and insurance and their characteristics in the form of expected value, variance, skewness and excess kurtosis. The claim distribution is of utmost importance in calculating premiums in non-life insurance. The actuary must estimate the frequency and severity of claims to determine the expected value of aggregate claim amount and the corresponding premium in accordance with the chosen premium principle. A proper selection of the claim distribution model is essential for the adequate pricing of insurance policies, as it directly affects the insurer's ability to cover potential losses. Insufficient premiums can lead to underestimation of technical reserves and insolvency of the insurer in the long run. We confirmed the importance of determining the appropriate probability distribution in claim data analysis through an example of premium calculation in motor vehicle (casco) insurance. The analysis carried out demonstrates that the insurance premium amount can vary considerably depending on the selected premium principles and coefficients. Therefore, it is essential to have experienced actuary who can select the distribution of losses, premium principles, and coefficients to ensure an adequate insurance premiums.