DOI: https://doi.org/10.62232/barp.9.2024.8570

About one method of pricing an insurance product

Erekle Khurodze

The insurance rate is the portion of the insurance limit the insurer must charge the insured in exchange for transferring the risk. It can be decomposed into two parts: the risk rate and the so-called loading. The first is responsible for covering future losses (which is, of course, probabilistic in nature), while the purpose of the second is to generate the funds necessary for business activities (such as salaries, profits, etc.).

The work of actuaries in tariffication typically involves offering the company a risk rate. If the insurance company already has experience in insuring the product in question (i.e., if it has accumulated certain statistical data), this information will be used when determining the new risk premium. Indeed, there are several classical methods for determining an "adequate" risk premium based on experience (an "adequate" risk premium is one that ensures coverage of future losses with a certain, reasonable reliability, and within which the insurance product will remain (in a certain sense) competitive in the market).

The above problem naturally leads to the concept of a break-even rate: for any fixed (completed) portfolio, the break-even rate is the rate at which the insurance company would go to zero profit in the given (existing, portfolio-specific) scenario. In the simple case, when all policies in a given portfolio have the same contribution (the duration of the policies is the same: $t_i \equiv t$; the portfolio includes the total duration of the policies), it takes the following form:

$$BE = \frac{\sum_i C_i}{\sum_i S_i} \tag{1}$$

where C_i are the loss amounts (which can be 0), while S_i are the insurance limits. A typical example of such a portfolio is a cohort — for instance, a set of policies "born" in a specific calendar year. Note also that in this case, we are referring to the corresponding rate for a policy with duration t.

In a more general case (which is more applicable when the portfolio is defined by fixing the calendar period, i.e., policies that were active for at least one day during some period P), we can consider the following (allocated) form:

$$BE = \frac{\sum_i C_i}{\sum_i S_i \frac{t_i}{\alpha}} \tag{2}$$

where C_i – total loss amount in the period P of the i -th policy, $t_i = |T_i \cap P|$, and T_i is the duration of the policy, while α is any duration of time (to which the calculated premium corresponds); or the following form:

$$BE = \frac{\sum_i \hat{C}_i}{\sum_i S_i \frac{T_i}{\alpha}} \tag{3}$$

where $\hat{C}_i$ – the total loss amount of the i-th policy, regardless of whether the loss occurred in the period P or not.

Both the numerator and the denominator of the fractions above are random in nature, and the BE calculated for any particular portfolio is only one manifestation of this randomness (only one realization of a random variable), by which to make a direct decision is, to put it mildly, naive.

The solution lies in studying the probabilistic nature of the BE (as of random variable). Its analysis using classical statistical methods, specifically by independent and identically distributed (i.i.d.) realizations, is usually not feasible, since using very old data in relation to a lively business is not advisable — there is a high chance that it no longer reflects the current reality. Additionally, considering many small (disjoint) portfolios is not practically profitable, as in this case, the dispersion of BE increases so much that the obtained results become practically useless.

Another approach may be to study all the random variables involved in the BE, determining the nature of their dependencies, and then use Monte Carlo methods to "artificially" generate a large amount of information, which can then be used to study the random nature of BE. This method is indeed actively used by insurance companies, although it is also associated with certain challenges: specifically, the fact that statistical dependencies between insurance indicators are often non-trivial and complex; the study and modeling of these dependencies (if it is possible) requires significant time and intellectual resources, which poses a challenge for insurance companies that need to solve many such problems within a reasonably short time. To illustrate this, and, most importantly, to discuss the method we have proposed, let us introduce the concept of an insurance (stochastic/marked point) process, which has the following form:

$$Z_n \equiv (K_n, r_n, S_n, H_n, C_n), \quad n \geq 1 \tag{4}$$

where K_n are the moments of appearance, or jumps, of the point process $((N_t)_{t\geq 0}$, that counts the insured policies up to moment t. In this context, we can use a homogeneous Poisson process with intensity λ (or intensity $\lambda(t)$ in the case of high seasonality).

$$P(N_{t+s} - N_t) = \frac{(s\lambda)^k e^{-s\lambda}}{k!} \quad \text{ან} \quad P(N_{t+s} - N_t) = \frac{\left(\int_t^{t+s} \lambda(y)dy\right)^k e^{-\int_t^{t+s} \lambda(y)dy}}{k!}$$

The remaining components are the marks of the point process:

r_n – the duration of the n-th appeared (insured) policy (typically constant);

S_n - the insurance limit of the n-th insured policy. For its modeling, various well-known families of continuous distributions are used, typically right-skewed ones;

H_n - the number of losses incurred by the n-th policy. For its modeling, Poisson, negative binomial, Bernoulli, and other well-known discrete distributions are typically used;

C_n - expresses the share of the total (overall) amount of losses of the n-th policy relative to the insurance limit. We can think of it as a mixture:

$$F_{C_n}(x) = \alpha_0 + \alpha_1 F_1(x) + \alpha_2 F_2(x) + \cdots$$

where α_i is the probability that the policy will suffer the i losses, and $F_i(x)$ is the distribution function of the share of the total amount of loss in the insurance limit, given that the policy has suffered the i losses. For example, if $H_n \sim Pois(\lambda)$:

$$\alpha_i = \frac{\lambda^i e^{-\lambda}}{k!}$$

To model the $F_i(x)$ distributions, Gamma, Weibull, Pareto, and other distributions are often used (although it should be noted that not with the direct form, since the support of C_n- is $[0,1]$, To avoid this inconvenience, the Beta distribution is sometimes used, although there is no inherent need for this (random variables of the type $C_n \equiv \min(Y_n, 1)$ are often useful)).

It is possible to consider α_i i.e., H_n, and $F_i(x)$, i.e., C_n, as independent of t, i.e., K_n, although it would not be natural to consider them as independent of S_n as well. Thus, more generally, we can write:

$$F_{C_n|S_n=s}(x) = \alpha_0(s) + \alpha_1(s)F_1\left(x, \lambda_1^{(1)}(s), \dots, \lambda_{m_1}^{(1)}(s)\right) + \alpha_2(s)F_2\left(x, \lambda_1^{(2)}(s), \dots, \lambda_{m_2}^{(2)}(s)\right) + \cdots$$

where $\lambda_1^{(i)}(s), \dots, \lambda_{m_i}^{(i)}(s)$ are the parameters of the distribution F_i, depending on s.

In the case where, for any S_n, the share of losses does not depend on the order of the losses — that is, if the shares of losses occurring first, second, etc., in the sum insured are independent and identically distributed random variables, with the distribution function of which is F, then:

$$F_{C_n|S_n=s}(x) = \sum_{i=0}^{\infty} \alpha_i(s)\left(\overline{*_{j=0}^{i-1} F}\right)(x)$$

where $*_{j=0}^{i} F = F * \dots * F, i \geq 1$, represents convolution of F, $i + 1$ times to itself, $*_{j=0}^{0} F \equiv 1$, and:

$$\bar{G}(x) \coloneqq \begin{cases} G(x), & x < 1 \\ 1, & x \geq 1 \end{cases}$$

Note that the given model is useful when using the (3) form of BE. In the case of using the (2) form, it is also essential to consider at what point in time the loss occurred. Therefore, in this case, the corresponding mark should be added to the (4) model. Let's start with a simple case; assume the following:

$$H_n \sim Bernoulli\big(p(S_n)\big)$$

That is, H_n is Bernoulli random variable (the probability of "success" depends on S_n), meaning that the loss either occurs (once) or does not occur. An example of this type of insurance is life insurance or so-called "insurance until the first loss". In such cases, model (4) can be easily generalized by adding one mark:

$$Z_n \equiv (K_n, r_n, S_n, H_n, C_n, K_n^c), \ \ n \geq 1 \tag{5}$$

where K_n^c is a random moment in the life period of the policy, that express the moment of loss occurrence, or it is 0 under the condition of no loss. Thus:

$$supp(K_n^c) = \{0\} \cup [K_n, K_n + r_n]$$

For example, if, in the loss condition, the moment of loss is uniformly distributed over the lifetime of the policy (which is a natural assumption under non-seasonality), then:

$$F_{K_n^c|S_n=s}(x) = \big(1 - p(S_n)\big) + p(S_n)F_{Unif([K_n,K_n+r_n])}(x)$$

where $F_{Unif([K_n,K_n+r_n])}(x)$ is the distribution function of a uniformly distributed random variable on the interval $[K_n, K_n + r_n]$.

If we do not assume that H_n is Bernoulli random variables, then model (5) can be generalized as follows:

$$Z_n \equiv \left(K_n, r_n, S_n, M_n = \left(K_{n,m}^c, C_{n,m}\right)_{m\geq 1}\right), \quad n \geq 1 \tag{6}$$

where M_n is a marked point process on the interval $[K_n, K_n + r_n]$ (which we can consider as a random element), where $K_{n,m}^c$ are the moments of occurrence of the point process $(N_\tau^c)_{\tau\in[K_n,K_n+r_n]}$, with intensity $\lambda(S_n)$, or, more generally, when the loss is seasonal, with intensity $\lambda(S_n, t)$ (for practical reasons, it might be more appropriate to consider an even more general form: $\lambda(S_n, t, t - K_n)$— when the loss depends on the period that has passed since the policy was started); And $C_{n,m}$ is the mark, which expresses the amount of the loss: the share of the loss amount in the insurance limit, the distribution of which also depends on S_n.

Obviously, in standard cases (when we are not dealing with insurance that involves the automatic restoration of the limit, thus the total amount loss should not exceed the insurance limit), instead of $C_{n,1}, C_{n,2}, \dots$ it would be more appropriate to consider the sequence $X_{n,1}$ ($\equiv C_{n,1}$), $X_{n,2}, X_{n,3}, \dots$ where:

$$X_{n,i} = \min\left(C_{n,i}, 1 - \sum_{j=1}^{i-1} X_j\right), \quad \forall i \geq 2$$

and consider the process

$$Z_n \equiv \left(K_n, r_n, S_n, M_n = \left(K_{n,m}^c, X_{n,m}\right)_{m\geq 1}\right), \quad n \geq 1 \tag{7}$$

Let us return to the task of estimating BE: obviously, if the probabilistic nature of the Z_n process is fully investigated, we can generate many of its "trajectories" through Monte Carlo simulations, generate many "point estimates" of BE, and study them. However, as we mentioned above and as became clear from the construction of the process, studying these relationships in detail is a rather difficult task. In essence (and in practice), we have only one realization of it: the real statistics of the company. Nevertheless, in non-parametric statistics, methods based on sub-selections are known, which we can also use in this case.

To do this, let us consider the stochastic process $\left(R_\beta\right)_{\beta\in\Gamma}$ derived from Z_n, where $\Gamma = \mathcal{B}([0,T])$ ($[0,T]$ is the time interval over which we observe the Z_n process (or the process $(N_t)_{t\geq 0}$ and its marks), and $\mathcal{B}([0,T])$ is the sigma-algebra of its Borel subsets; for practical purposes, we could obviously just take the set of open subsets). This process can be called the $\boldsymbol{\alpha}$ **risk-rate** process and defined as follows: for model (5):

$$R_\beta = \frac{\sum_{K_n^c\in\beta} C_n S_n}{\sum_{K_n\in[0,T]} S_n \dfrac{|\beta \cap [K_n,\ K_n + r_n]|}{\alpha}} \tag{8}$$

and for model (7):

$$R_\beta = \frac{\sum_{K_n\in[0,T]} \left(\sum_{K_{n,m}^c\in\beta} X_{n.m} S_n\right)}{\sum_{K_n\in[0,T]} S_n \dfrac{|\beta \cap [K_n,\ K_n + r_n]|}{\alpha}} \tag{9}$$

From the one "trajectory" of Z_n, we can obtain one "trajectory" of $\left(R_\beta\right)_{\beta\in\Gamma}$; and after fixing the length of time $\hat{\alpha}$, the set:

$$\{R_\beta\colon\ |\beta| = \hat{\alpha}\}.$$

Obviously, the latter is an infinite set, but for practical purposes, by means of sub-selections, for any fixed n_0, we can obtain (one of) the following set(s):

$$\mathbf{BE}_{n_0,\hat{\alpha}} \subset \{R_\beta\colon\ |\beta| = \hat{\alpha}\}, \quad \text{where } \ card(\mathbf{BE}_{n_0,\hat{\alpha}}) = n_0,$$

Through this, we can estimate the distribution of BE (the empirical results of this procedure for a specific example are discussed below).

With the inspired of Bradley Efron's bootstrap idea, we considered a more general model:

$$\left(\tilde{R}_\beta\right)_{\beta\in\Gamma_1}, \text{ where } \ \Gamma_1 = \left\{\prod\nolimits_{i=1}^{\kappa} \beta_i :\ \forall i\colon \beta_i \in \mathcal{B}([0,T])\ \&\ card(\kappa) < \aleph_0\right\}$$

and (for model (5)):

$$\tilde{R}_\beta \equiv \tilde{R}_{\prod_{i=1}^{\kappa}\beta_i} = \frac{\sum_{i=1}^{\kappa}\left(\sum_{K_n^c\in\beta_i} C_n S_n\right)}{\sum_{i=1}^{\kappa}\left(\sum_{K_n\in[0,T]} S_n \dfrac{|\beta_i \cap [K_n,\ K_n + r_n]|}{\alpha}\right)} \tag{10}$$

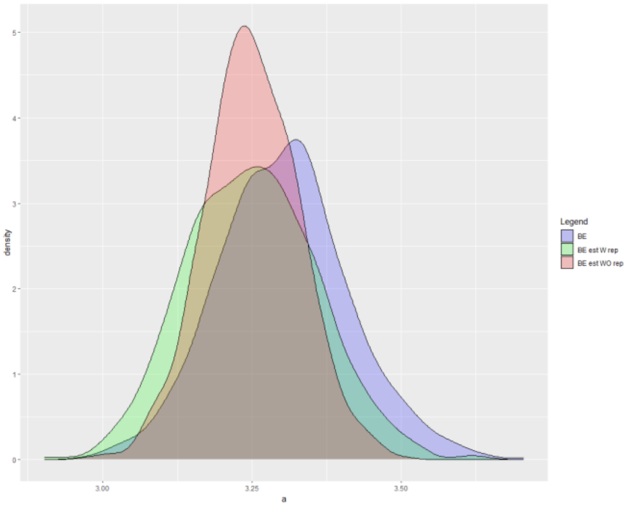
or (for model (7)):

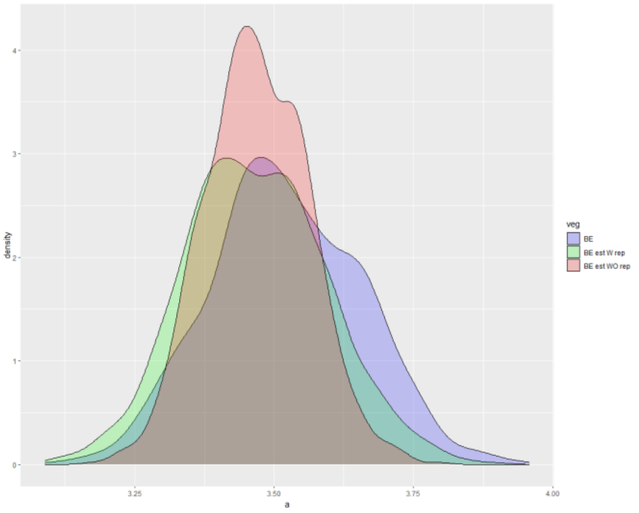
$$\tilde{R}_\beta = \frac{\sum_{K_n\in[0,T]}\left(\sum_{i=1}^{\kappa}\left(\sum_{K_{n,m}^c\in\beta_i} X_{n.m} S_n\right)\right)}{\sum_{i=1}^{\kappa}\left(\sum_{K_n\in[0,T]} S_n \dfrac{|\beta_i \cap [K_n,\ K_n + r_n]|}{\alpha}\right)} \tag{11}$$

In this case as well, for fixed α and n_0 , we can obtain:

$$\overline{\mathbf{BE}}_{n_0,\hat{\alpha}} \subset \left\{\tilde{R}_\beta \equiv \tilde{R}_{\prod_{i=1}^{\kappa}\beta_i}\colon\ \sum\nolimits_{i=1}^{\kappa} \beta_i = \hat{\alpha}\right\}, \quad \text{სადაც}\ card(\overline{\mathbf{BE}}_{n_0,\hat{\alpha}}) = n_0.$$

Below are the simulation results (using the density Kernel estimates), where the results of $\mathbf{BE}_{n_0,\hat{\alpha}}$ and $\overline{\mathbf{BE}}_{n_0,\hat{\alpha}}$ are compared to the BE distribution for the time length $\hat{\alpha}$ for two examples. In the first (left), it is assumed that H_n and C_n are independent of S_n, while in the second (right), they are not (we will not provide the technical details of the dependence here):





The blue figure is bounded by the density estimate of the BE distribution, the green figure is bounded by the density estimate obtained from $\overline{\mathbf{BE}}_{n_0,\hat{\alpha}}$, and the red figure is bounded by $\mathbf{BE}_{n_0,\hat{\alpha}}$. As we can see, the estimations with the so-called "subsampling with replacement" ($\overline{\mathbf{BE}}_{n_0,\hat{\alpha}}$) is better than those with the so-called "subsampling without replacement" ($\mathbf{BE}_{n_0,\hat{\alpha}}$).

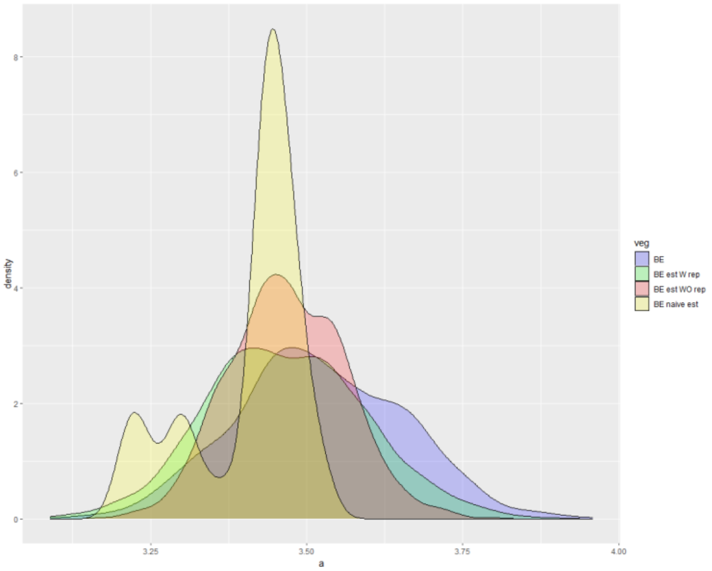
Due to simplicity and, at first glance, naturalness, it may seem acceptable to consider the following special case of the above processes:

$$\left(\hat{R}_t\right)_{0\leq t<T-\hat{\alpha}}, \quad \text{where } \hat{R}_t = R_{[t,\, t+\hat{\alpha}]}$$

and obtaining an estimate through it:

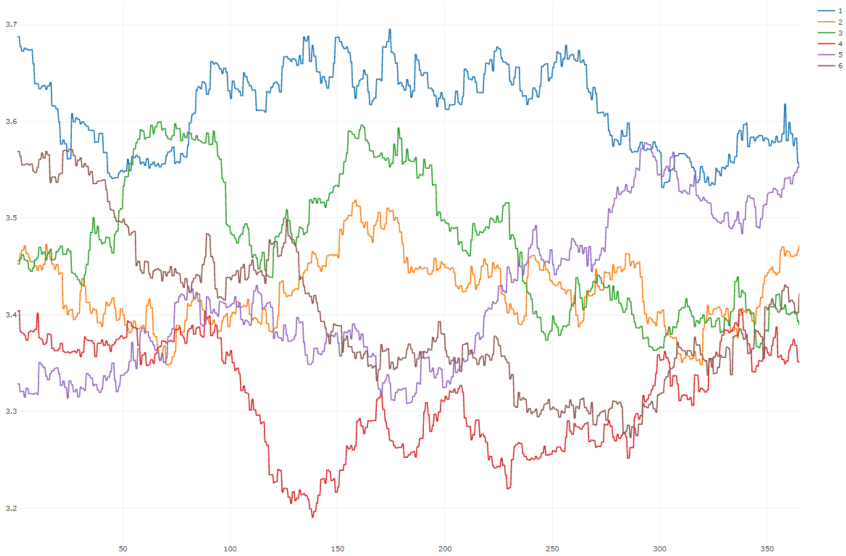
$$\overline{\overline{\mathbf{BE}}}_{n_0,\hat{\alpha}} \subset \{\hat{R}_t : 0 \leq t < T - \hat{\alpha}\}$$

Let us compare the result obtained from (12)-(13) with the results obtained above:

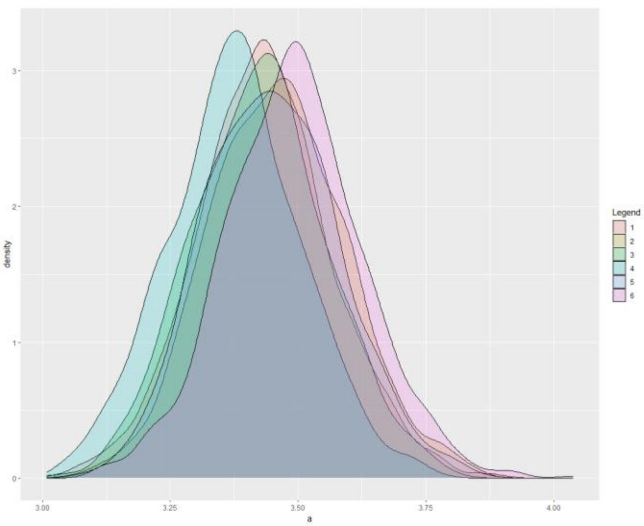


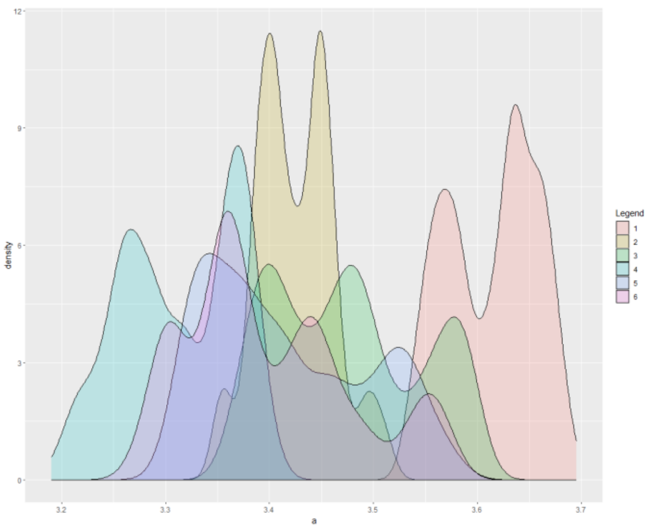
The result obtained from (12)-(13) is highlighted in yellow. As we can see, the quality of estimation is much lower than the estimates obtained above.

Through simulations, we can empirically observe one more thing: during the estimation, how we are momentarily dependent on the one trajectory that is "in our hands." Consider the six trajectories of Z_n and the trajectories of $\tilde{R}_\beta$ and $\hat{R}_t$ obtained from it (the image below shows the trajectories of $\hat{R}_t$).



For which the density estimates are as follows:





As we can see, $\left(\hat{R}_t\right)_{0\leq t<T-\hat{\alpha}}$, despite its naturalness, does not provide reliable estimates: the estimate depends significantly on the trajectory. This cannot be said about the estimate obtained with $\left(\tilde{R}_\beta\right)_{\beta\in\Gamma_1}$, which seems quite reliable for practical purposes.

References:

(1) Vladimir I. Rotar - Actuarial Models: The Mathematics of Insurance;
(2) B. Efron - Bootstrap Methods: Another Look at the Jackknife;
(3) Stanisław Węglarczyk - Kernel density estimation and its application;
(4) Alan F. Karr – Point processes and their statistical inference.