**Chapter 18.**

# UNCERTAINTY ANALYSIS AND RISK MODELING IN INSURANCE

There are various definitions of risks, depending on the used approach, context and the given aspect. It is considered by some authors that risk and uncertainty are two completely independent concepts, while others see these terms as highly interdependent. The notion of uncertainty implies that there is a possible spectrum of events that can happen in the future, however without indications that they will really happen or have a relative effect. On the other hand, the term risk implies the possibility to list events that may occur in the future, and that the likelihood of realization can be determined for the each one of them. The uncertainty is also described as the absence of information in a decision-making process and implies the assessment of a particular situation, alternative solutions, possible results and consequences, etc. In extreme situations, uncertainty can be characterized as the lack of information or knowledge about a particular problem or in the decision-making process. Risk assessment is a complex task that requires a great deal of responsibility, so the actuary needs to be familiar with the application of the statistical theory of credibility and its basic methods. In this study basic credibility procedures, with appropriate calculations, are presented. Actuarial science requires a combination of academic rigor and business practice. Actuaries rarely use the theory of credibility in a purely statistical way. The subset of the observed population has characteristics that are not completely determined, so the actuarial assessment is necessary in order to determine the possible purpose of these characteristics. In most cases, the effects of population characteristics are known to some extent, but there is insufficient data to eliminate the need for a particular estimation in the selection of probability distribution and parameters. Numerous examples from practice indicate that there is often uncertainty about input data necessary for making certain decisions. Decisions are often made on the basis of experience, intuition, subjective assessment of paramatars that they appear in these situations. Fuzzy mathematical modeling is used in situations of uncertainty, uncertainty, and subjective evaluation.

## 1. CREDIBILITY THEORY

The credibility is an estimate of the predictive value in a certain situation that the actuary assigns to a particular set of data. The credibility procedure is a process that involves the estimation of the insurance company's experience for potential use in determining assumptions without references to other data or identifying the related experience, as well as the selection and application of the method for combining the relevant experience and the experience of the company. The credibility procedure is used to improve the estimation of a parameter in a given task. Credibility can be used to determine the price, calculate the premium rate, as well as to determine the future premium rate based on experience and reservation, etc.[461] The application of the credibility procedure requires taking into account the characteristics of both the experience of the insurance company and the relevant experience (experience similar to the experience of the company). Also, the prediction value of the latest experience should be considered in comparison with the experience from previous periods. The actuary should perform an expert assessment and carefully select and use the relevant experience. Characteristics to be considered include demographics, coverage, frequency, or other risk characteristics that can be determined, for which the actuary expects to be similar to the company's experience.

The use of the credibility method is not always a precise mathematical process. For example, in some situations, an acceptable procedure for combining the experience of a company with relevant experience can be based on the fact that the actuary grants a full, partial or no experience of a company without the use of a rigorous mathematical model. The actuary should use a professional judgment in the selection, development or the application of credibility procedures. During the decision process, the actuary should take into account the extent to which the experience of company is included in the relevant experience. If the experience of a company is an essential part of the relevant experience, the actuary should, based on a professional judgment, decide whether and how to use this relevant experience. Also, the homogeneity of these experiences should be taken into account, whereby segments that are not typical representatives of the experience as a whole can be excluded, thus obtaining a better predictive value. The actuary should also take into account the balance between the data homogeneity and the size of the data set.

## 1.1. The current practice in applying credibility concepts

There are different approaches in the implementation of credibility procedures. In some cases, the approach is based on estimation, while others deploy mathematical models. Some of the mathematical credibility procedures are discussed below. The most commonly used credibility procedures are based on

[461] Actuarial Standards Board (2011). Credibility Procedures Applicable to Accident and Health, Group Term Life, and Property/Casualty Coverages. *Actuarial Standard of Practice*, No. 25., Washington, DC: ASB, p. 2.

assumptions regarding the shape of the basic probability distribution. Based on this probability distribution function, the corresponding number of claims, as well as the amount of the premium, etc. is calculated, so that the probability of the claims occuring is within the specified percentage of the expected value. In the limited fluctuation credibility approach, it is assumed that claims follow normal distribution. Within this approach, partial credibility is assigned to the experience of the insurance company based on the square root of the ratio of actual claims to the standard of total credibility.

Empirical credibility procedures measure the statistical relationships of the subject experience to its mean and to comparable experience of prior periods, without reference to the underlying distribution.

Bayesian analysis combines current observations with a priori information in order to get the most accurate estimates, while the credibility theory directly depends on the existence of a priori information that could be weighted by current observations. In certain situations, the Bayes analysis formula exactly matches those of Buhlmann credibility assessment, due to the fact that Bayes estimate represents the linear weighting of the current and a priori information that is further weighted with Z and (1-Z), where Z stands for Buhlmann credibility. One example of Bayesian credibility is high precision credibility, also reffered to as Linear Bayes credibility or Buhlmann credibility.

The latest credibility methods include the estimation of credibility in generalized linear models or other multivariate modeling techniques. The most common forms of these models are often referred to as generalized linear mixed models, hierarchical models and models of mixed effects. In these models, credibility can be estimated on the basis of the statistical significance of the estimation of a parameter, the impact of the model on a split data set, or the consistency of any of these measures.

## 1.2. The application of credibility theory

The difference between statisticians and actuaries lies in the fact that, for example, a statistician will be 95% sure that a certain estimate is correct, which rarely occurs in practice. On the other hand, the task of the actuary is to determine what is to be done when there is no 95% confidence, but for example 45% or even when there is no knowledge of the confidence level. It is the task of the actuary to point out the way in which the existing confidence in the available data, other relevant information as well as certain important business aspects could be considered at the same time.

The credibility theory provides tools for working with random variables (data) used to predict future events or costs. For example, the insurance company uses data on past losses of insured (claims) to estimate future expenses for the provision of insurance coverage. However, insurance losses arise from random occurrences. The average annual cost of claims in the past few years may be a poor estimate of the costs for the next year. The expected accuracy of this estimate is the function of claims variance. These data by themselves are not acceptable for calculating the insurance rate.

Instead of relying solely on recent observations, better estimates can be obtained by combining these data with other information. For example, let's assume that recent experience suggests that workers should be charged a premium rate of 5 EUR (per 100 euros payroll) for employee insurance. Suppose the current rate is 10 EUR. What should be the new rate? Is it 5, 10, or a value between them? Credibility is used to weight these two estimates together, as follows:

$$\text{Estimated} = \text{Z} \times \left[\text{Observation}\right] + \left(1\text{-Z}\right) \times \left[\text{Other informations}\right],\ 0 < \text{Z} < 1,$$

where Z credibility is assigned to observation, and 1-Z is referred to as the complement of credibility.

The equation implies that if the set of observed data is large, the parameter will not vary greatly from one period to the next, and then Z will be closer to 1. On the other hand, if the observation consists of limited data, then Z will be closer to zero and more weight will be given to data from other sources. Credibility implies a linear assessment of the real expectations made as a result of a compromise between observation and the previous hypothesis. Thus, for the previous example, the premium rate would be Z x [5] + (1-Z) x [10].

Let's consider the following example. In a large population of car drivers, the average driver makes a claim every five years, i.e. the annual claim frequency is 0,25 per year. The result of a performed random selection is a driver who had three claims over the last five years with a claim frequency of 0,55 per year. An estimate of the expected future frequency for this driver is needed. If we had no information about the driver, except that he belongs to the observed population, the frequency would have been 0,25. However, we know that the claim frequency for the observed driver was 0,55. Is this our estimate of his future frequency of claims? There is a correlation between the previous claim frequency and the future claim frequency, but this correlation is not perfect. Due to the randomness of accidents, even the good drivers with low expected claim frequencies can be involved in one. On the other hand, drivers with poor driving skills can file no claims for couple of years..

It can be assumed that the right frequency is a value between 0,25 and 0,55. Expected value of future frequency of driver claims is equal to: $Z \cdot 0{,}55 + (1 - Z) \cdot 0{,}25$.

Next, it is necessary to determine Z, or how much credibility should be assigned to driver information. In an attempt to forecast future claims, the actuary does not know the value of the distribution parameters. Also, there are many possible and different values that can be best estimates of the basic parameters. One of these values is the recent historical experience of the company. Other values are reported in various industrial studies of insurers considered to be similar to the company for which the actuary is trying to predict the cost of the claims. How can the actuary select the best estimate that will be used in the forecast and how can they justify their selection process?

The credibility theory attempts to solve the problem with a compromise solution: instead of choosing one or the other best estimate, a value that is a linear combination or a weighted average of the best estimate is chosen. The credibility theory is used in the following procedure: as a preliminary step, from a variety of external or industrial assessments, the one that is rated as closest to the characteristics, considering the line of business, for which the actuary attempts to predict the future cost is chosen. It the next step combines the estimate obtained from the company's experience with the best outside, using the appropriate weights, in order to predict the future cost. Weight factors are determined as a function of the number of insured companies that the company had in the most recent historical period. The basic idea is to give more weight to the company's historical experience.

Using credibility theory, an actuary can anticipate future values in the following way: begin with an estimate of the cost from the company's latest historical experience - $\mu_{comp}$ and estimating costs derived from the most appropriate external source - $\mu_{ind}$. Then the values of $N_0$ i $N_{fcr}$ are determined. $N_0$ is the minimum sample size. If the company's costs come from a sample size that is smaller than $N_0$, then the company's experience is ignored and industry estimates are used. Similarly, $N_{fcr}$ is the size of the sample needed for full credibility. If the cost of damages to the company originates from a sample size that is greater than $N_{fcr}$, the company's experience is used and the industry's assessment is ignored.

Denote the number of insureds in the most recent historical period of the company with $N_{comp}$, and the forecasted claim cost with $N_{forecast}$. According to the credibility theory, it is estimated $N_{forecast}$ in the following way:[462]

$$If\ N_{comp} \leq N_0, \text{ then } C_{forecast} = \mu_{ind}$$
$$If\ N_{comp} \geq N_{fcr}, \text{ then } C_{forecast} = \mu_{comp}$$
$$If\ N_{comp} > N_0 \text{ and } N_{comp} < N_{fcr}, \text{ then } C_{forecast} = Z\,\mu_{comp} + (1 - Z)\ \mu_{ind}.$$

There are two methods for determining the weight in the above formula. According to the limited fluctuation method, $Z$ is determined as follows:

$$Z = \min\left(\sqrt{\frac{N_{comp}}{N_{fcr}}}, 1\right)$$

According to the second method, models of a greatest accuracy credibility are used. The basic element of this approach is that the concept of probability should be objective and free from the influence of any subjective factors, with the simultaneous possibility of practical application in experiments. According to this understanding, the probability can be defined and applied in situations that can be repeated over and over again under identical conditions. $Z$ is determined as follows:

$$Z = \frac{N_{comp}}{N_{comp} + K}$$

$K$ is another parameter to be determined. In order to apply these formulas in practice, the actuaries need estimates of the values of $N_0$, $N_{fcr}$, $Z$ and $K$. However, there are no simple and generally accepted ways of estimating these parameters, which limits the practical application of some of the credibility formulas. There are some problems in evaluating the value of the parameters. Actuarial literature does not provide any method of assessment of $N_0$, which is generally accepted and based on a scientific basis. The limited fluctuation approach gives the following formula for performing $N_{fcr}$:

$$N_{fcr} = \frac{z_{(1-p)}^2 \sigma^2}{h^2 \mu^2}$$

---

[462] American Academy of Actuaries (2008). *Credibility Practice Note*. Washington, DC: American Academy of Actuaries, p. 9.

However, for the application of limited credibility, it is necessary to define four parameters: *p, h,* $\mu$ i $\sigma^2$. Parameters *p* and *h* are the level of reliability and marginal error. By definition, each individual can choose any value for these parameters. The fact is that in practice the most commonly used 95% level of reliability and 5% for marginal error, does not provide scientific foundation. The actuary might want to use the other values for these parameters. As a result, for the same insurance, an actuary may consider, for example, 400 for the sample size that produces complete credibility for the experiential data. The other actuary may take into account any number of samples less than 1100 that is not sufficient the complete credibility.

Parameters $\mu$ and $\sigma^2$ are also unknown. If the actuary succeeds in determining $\mu$ (the value of a certain size per month per member), it is not necessary to determine the $N_{fcr}$. In actuarial literature *K* is defined as the ratio $\nu / \alpha$, where $\nu$ is the expected process variance and $\alpha$ is defined as the variance of the hypothetical means. There are no good and acceptable estimates of these values. In fact, there are no estimates of these parameters at all. Without good estimates of these two parameters, the estimation of *K*, and therefore the assessment of *Z* becomes problematic.

## 2. LIMITED FLUCTUATION APPROACH

We need the following data to calculate the credibility factor Z, in accordance with the method:[463]

1. *n* opserved insured persons at a certain point, ie part of the year $t_i$ ,
2. Sum insured $B_i$,
3. $d_i = 0$ if insured person did not die and $d_i = 1$ if the person did die,
4. Let us denote by $q_i$ the real mortality rate and $q_i^s$ the mortality rate from the Mortality Table and suppose that $q_i = mq_i^s$, where m is the real racio.[464]
5. *A* is the real, and *E* is the expected value.

We can now determine the following values:

---

[463] Klugman S., Rhodes, T., Purushotham, M., Gill, S. & MIB Solutions. (2009). *Credibility Theory Practices*. Schaumburg, IL: Society of Actuaries, p. I.4.

[464] It is possible to create models where the multiplier depends on a particular category (such as age or duration). However, the relationship must be explicit and the formulas become much more complex.

$$A = \sum_{i=1}^{n} B_i d_i$$
$$E = \sum_{i=1}^{n} B_i f_i q_i^s$$
$$\hat{m} = A / E$$

The preceding formulas refer to amounts, if we work with numbers, simply substitute the sum of the insurance with the numbers (number of claims, number of lapses etc.).

Further, assuming $q_i = mq_i^s$, we get the following formulas:

$$\mu = E(\hat{m}) = \frac{\sum_{i=1}^{n} B_i E(d_i)}{E} = \frac{\sum_{i=1}^{n} B_i f_i q_i}{E} = m$$
$$\sigma^2 = Var(\hat{m}) = \frac{\sum_{i=1}^{n} B_i^2 Var(d_i)}{E^2} = \frac{\sum_{i=1}^{n} B_i^2 f_i q_i (1 - f_i q_i)}{E^2}$$

$$= \frac{\sum_{i=1}^{n} B_i^2 f_i m_d q_i^s (1 - f_i m_d q_i^s)}{E^2} \quad (1)$$

To determine the variation for numbers, the equation has following form:

$$\sigma^2 = Var(\hat{m}) = \frac{\sum_{i=1}^{n} Var(d_i)}{E^2} = \frac{\sum_{i=1}^{n} f_i q_i (1 - f_i q_i)}{E^2} = \frac{m}{E} = \frac{A}{E^2} \quad (2)$$

The credibility formula now becomes $Z\hat{m} + (1-Z)\alpha$, where $Z$ and $\alpha$ need to be determined, whereby the method does not specify how these parameters should be determined.

For $Z = 1$, we have full credibility if $Pr(|\hat{m} - m| \le rm) \ge p$, and it represents a relative error. The method also does not specify how to determine $r$ and $p$. If this condition is not met, then Z is chosen in such a way to reduce the variance of the credibility assessment to the value where it has the desired accuracy.

It is commonly to use a normal approximation in order to estimate the probability and then the credibility factor [465]

$$Z = \min\left(\frac{r\hat{m}}{z\hat{\sigma}},1\right) \tag{3}$$

The value of $\alpha$ is usually taken as a ratio that would be used if there are no data (ie, assigning is given zero credibility). If we assign 1, it implies that a standard table, or average mortality in the observed companies, is used. The choice of $\alpha$ can have a significant impact on the final result.

Example 1: we want to estimate the mortality rate q, for the group of insured persons aged 55 to 65 years, with a one-year experience for 1.000 insured persons with data as in Table 1.

*Table 1. Data for Example 1.*

| No. of policy | Sum insured | No. of dead |
|---|---|---|
| 120 | 8.000 | 5 |
| 270 | 27.000 | 9 |
| 460 | 45.000 | 7 |
| 150 | 80.000 | 2 |

Estimates of the mortality rate are calculated using the formula:

$$\hat{q} = \frac{\sum_{i=1}^{1000} B_i d_i}{\sum_{i=1}^{1000} B_i} \tag{4}$$

The method does not provide us with guidelines in setting parameters to make the assessment complete credibility, but leaves a free choice. We will assume that the probability of a relative error is less than or equal to 5% at least 90%, and that Z has a standardized normal distribution, that is, we verify that:

$\Pr\left(\frac{|\hat{q}-q|}{q} \le 0,05\right)$ at least 90%.

Assumptions:
a) Central Limit Theorem can be applied
b) The amounts of the insurance sum are not random, and

[465] Corresponding quantile from the normal distribution, based on the chosen value of *p* (for example, for the confidence interval *p* = 90% is *z* = 1.64485).

c) The insured pearsons are mutually independent and have the same value q.

Then:

$$\Pr\left(-0,05q \leq \hat{q} - q \leq 0,05q\right) = \Pr\left(\frac{-0,05q}{\sigma} \leq Z \leq \frac{0,05q}{\sigma}\right)$$

*Table 2.*

| No. of policy | Sum insured | No. of dead | $B_i d_i$ | $B_i$ |
|---|---|---|---|---|
| 120 | 8.000 | 5 | 40.000,00 | 960.000,00 |
| 270 | 27.000 | 9 | 243.000,00 | 7.290.000,00 |
| 460 | 45.000 | 7 | 315.000,00 | 20.700.000,00 |
| 150 | 80.000 | 2 | 160.000,00 | 12.000.000,00 |
| 1000 | 160.000 | 23 | 758.000,00 | 40.950.000,00 |

From where, using formula (4) that is:

$\hat{q} = 0,0185,\ \ \sigma = 0,005$ and $\Pr\left(-0,2035 \leq Z \leq 0,2035\right) = 0,1586,$

and that is less than 0.9 so the assessment does not have full credibility. Since the full credibility has not been reached, the weight must be determined, which is called the partial credibility factor. The most common method used is the rule of a square root. There are several justifications for this approach, but they all have flaws.

Firstly, it is necessary to determine the minimum exposure required for full credibility, so the weight is the square root of the actual exposure and the minimum established exposure. The same value is obtained if we determine the standard normal value for the desired probability (confidence interval), and then to determine the ratio of the real value and this value. For example, the partial credibility factor is 0,2022 / 1,64485 = 0,1229.

This suggests that a weight of 12.29% should be given to the observed mortality rate (relative frequency). The method does not determine what to do with the rest of the weight, on which value to apply.

The next example shows credibility calculation for lapses in life insurance. Table 3 contains data from 6 insurance companies, the number and the amount of real and expected lapses. For the illustration, a mix of large, medium and small insurance companies are selected.

*Table 3. Data for six companies*

| | No. lapses | Amount of lapses | Expected number of lapses | Expected amount of lapses |
|---|---|---|---|---|
| C1 | 12.012,00 | 789.113.322,00 | 9.620,54 | 590.962.948,06 |
| C2 | 1.066,00 | 69.840.534,00 | 1.170,55 | 115.278.432,01 |
| C3 | 3.874,00 | 880.443.656,00 | 3.331,02 | 670.024.330,00 |
| C4 | 1.230,00 | 56.965.533,00 | 890,00 | 33.593.401,80 |
| C5 | 1.803,00 | 291.423.661,00 | 1.902,12 | 340.221.381,20 |
| C6 | 4,00 | 1.250.400,00 | 2,33 | 688.105,65 |
| | 19.989,00 | 2.089.037.106,00 | 16.916,56 | 1.750.768.598,72 |

In the first step, the A / E ratio is calculated for the number and amount of the lapsed contracts, as well as the variance according to equations (1) and (2). The results are shown in Table 4.

*Table 4. A/E and variance calculation*

| | A/E lapse (number) | A/E lapse (amount) | Variance (number) | Variance (amount) |
|---|---|---|---|---|
| C1 | 1,25 | 1,34 | 0,00014 | 0,00022 |
| C2 | 0,91 | 0,61 | 0,00076 | 0,00525 |
| C3 | 1,16 | 1,31 | 0,00035 | 0,00126 |
| C4 | 1,38 | 1,70 | 0,00224 | 0,16792 |
| C5 | 0,95 | 0,86 | 0,00049 | 0,00219 |
| C6 | 1,72 | 1,82 | 0,53334 | 1,60133 |
| | **1,18** | **1,19** | | |

*Table 5. Lapse Results by Policy for Limited Fluctuation*

| | Overall A/E Ratio | Z | | Comp. A/E Ratio | Number of Lapses | Estimated A/E Ratio | |
|---|---|---|---|---|---|---|---|
| | | (p=0,9) | (p=0,95) | | | (p=0,9) | (p=0,95) |
| C1 | 118,2% | 1,000 | 1,000 | 124,9% | 12.012 | 124,9% | 124,9% |
| C2 | 118,2% | 1,000 | 1,000 | 91,1% | 1.066 | 91,1% | 91,1% |
| C3 | 118,2% | 1,000 | 1,000 | 116,3% | 3.874 | 116,3% | 116,3% |
| C4 | 118,2% | 1,000 | 0,637 | 138,2% | 1.230 | 138,2% | 130,9% |
| C5 | 118,2% | 1,000 | 1,000 | 94,8% | 1.803 | 94,8% | 94,8% |
| C6 | 118,2% | 0,098 | 0,041 | 171,7% | 4 | 123,4% | 120,4% |

In the Table 5 calculations were performed with confidence interval $p$ = 90% and $p$ = 95% by Policy and by Amount, whereby estimated A/E ratio is:

$$Z \times [\text{Company A/E ratio}] + (1\text{-}Z) \times [\text{Overall A/E ratio}],\ 0 < Z < 1,$$

*Table 6. Lapse Results by Amount for Limited Fluctuation*

| | Overall A/E Ratio | Z | | Comp. A/E Ratio | Number of Lapses | Fluctuation A/E Ratio | |
|---|---|---|---|---|---|---|---|
| | | (p=0,9) | (p=0,95) | | | (p=0,9) | (p=0,95) |
| C1 | 119,3% | 1,000 | 1,000 | 133,5% | 12.012 | 133,5% | 133,5% |
| C2 | 119,3% | 1,000 | 0,420 | 60,6% | 1.066 | 60,6% | 94,6% |
| C3 | 119,3% | 1,000 | 0,859 | 131,4% | 3.874 | 131,4% | 129,7% |
| C4 | 119,3% | 0,177 | 0,074 | 169,6% | 1.230 | 128,2% | 123,1% |
| C5 | 119,3% | 1,000 | 0,650 | 85,7% | 1.803 | 85,7% | 97,4% |
| C6 | 119,3% | 0,057 | 0,024 | 181,7% | 4 | 122,9% | 120,8% |

## 2.1. Summary of limited fluctuation credibility

The advantage of the model is in the simple implementation and comprehensibility. Limited fluctuation credibility approach is suitable for determining experience rating where there is a default premium. In certain cases, no assessment is required to determine whether there is full credibility or to calculate a partial credibility factor.

On the other hand, the lack of a model is only to reflect the variability of the sampled data. Parameters, $\alpha$ , $r$ and $p$ are arbitrarily chosen and the accuracy of the parameter $\alpha$ is not included in the calculation of Z.

# 3. GREATEST ACCURACY CREDIBILITY METHOD

Greatest Accuracy Credibility is also referred to as Bayesian credibility, linear Bayesian credibility and Buhlmann credibility. The underlying element of this approach is that the concept of probability should be objective and free from the influence of any subjective factors, with the simultaneous possibility of practical application in experiments. According to this method, the probability can be defined and applied in situations that can be repeated over and over again under identical conditions.

All variations of the method assume that there are more than one entity that has their own probability distribution. Another assumption is that these probability distributions are distributed among the entities in accordance with the second distribution of probability. The goal is then to use this information to estimate the distribution of the probability (or key parameters of that distribution) for one or all entities.

One of the differences in the methods is to use the total variance, which also affects the result. The total variance for observations is the sum over all entities and the two variation obtained from different sources:

1. For each entity, the variation of entities' observations about that company's mean.
2. Variation between the mean of each entity and the total mean.

There a coulpe of limitations in this approach. First, the appropriate distributions must be identified, and then their parameters. Also, the implementation of the Bayesian' theorem is not always simple and depends on the distribution. In the case of normal distribution, there is no difficulty in obtaining the Bayesian solution. It is:

$$Z\overline{x}+\left(1-Z\right)\mu \tag{5}$$

$$Z=\frac{n}{n+\nu/\alpha} \tag{6}$$

If there are more observations (higher n), the sample mean has higher credibility (although it is never complete credibility). If $\nu$ is high then the observations are very variable, which means that the data is less credible, which is what the formula indicates. Let's assume that $\alpha$ is high. Then the entities are very different from each other. Therefore, each of them can be quite large or small and therefore should not move towards the middle. This implies greater credibility. On the other hand, if $\alpha=0$, each entity would have its own mean and there is no reason to grant credentials to data, i.e. each entity is average. Another way of looking at this is that $\alpha$ refers to quality of $\mu$ in the same way as it $\nu$ refers to the quality of the sample mean.

Suppose we are looking at a group in which 2000 insured lives and 23 deaths and another group with 3,000 live and 67 deaths. The empirical Bayes calculation approach would be as follows in Table 7.[466]

The mean is $\mu=90/5000=0,018$.

The variance is determined by the variance estimation for each group, and then weighted by their sample size with a small correction to obtain unbiased estimate. The result is $\nu=0,01765$.

466 Klugman, S. A., Panjer, H. H., & Willmot, G. E. (2004). The *Loss Models: From Data to Decisions.* John Wiley & Sons, p. 595.

*Table 7. Empirical Bayes calculation approach*

| Lives | Deaths | q | 1-q | Var |
|---|---|---|---|---|
| 2000 | 23 | 0,0115 | 0,9885 | 22,7355 |
| 3000 | 67 | 0,0223 | 0,9777 | 65,5037 |
| 5000 | 90 | | | 88,2392 |
| $\mu =$ | 0,018 | | $\nu =$ | 0,01765 |

Futher calculation is:

$$\hat{\alpha} = \frac{2000\left(0,0115-0,018\right)^2 + 3000\left(0,02233-0,018\right)^2 - 0,0176}{5000 - \dfrac{2000^2+3000^2}{5000}}$$

$\hat{\alpha} = 0,00005132.$

$Z$ for the first group is obtained by applying the equation (6):

$$Z = \frac{2000}{2000 + 0,01765 / 0,00005132} = 0,8532,$$

further using formula (5) credibility estimate is 0,0124.

In the Bayesian model an initial or prior distribution is based on the past data, professional experience, and/or opinion. Observed results are then used to formulate a predictive or posterior distribution.

In further research we will present Bühlmann empirical Bayesian Method based on a linear Bayesian model that uses only first two moments of distribution. The method is empirical, since the moments of prior distribution are based on past data.

$$E\left(\hat{\mu}\right) = \mu \text{ and } \hat{\mu} = \frac{\sum_{i=1}^{n} A_i}{\sum_{i=1}^{n} E_i},$$

Equation for Variance is:

$$\hat{\sigma}^2 = \frac{\sum_{i=1}^{n} E_i \left(\hat{m}_i - \hat{\mu}\right)^2 - \hat{\mu}\left(\sum_{i=1}^{n} \frac{B_i}{E_i} - \frac{1}{\sum_{i=1}^{n} E_i}\sum_{i=1}^{n} B_i\right) + \hat{\mu}^2\left(\sum_{i=1}^{n} \frac{C_i}{E_i} - \frac{1}{\sum_{i=1}^{n} E_i}\sum_{i=1}^{n} C_i\right)}{\sum_{i=1}^{n} E_i - \frac{1}{\sum_{i=1}^{n} E_i}\sum_{i=1}^{n} E_i^2 - \sum_{i=1}^{n} \frac{C_i}{E_i} + \frac{1}{\sum_{i=1}^{n} E_i}\sum_{i=1}^{n} C_i}$$

The credibility factor is derived from the following formula:

$$Z = \frac{Ei}{E_i + \frac{\mu}{\sigma^2 E_i} B_i - \frac{\mu^2 + \sigma^2}{\sigma^2 E_i} C_i} \tag{7}$$

where:

$B_i$ - sum of the expected contracts lapsed for entity, and

$C_i$ - sum of the square of the expected contracts lapsed for each policy of entity $i$.

Estimated A/E ratio is:

$$\mathrm{Z} \times \left[entity\ A/E\ ratio\right] + \left(1\text{-}\mathrm{Z}\right) \times \left[Overall\ A/E\ ratio\right].$$

*Table 8. Number of lapses by policy*

| | Number of lapses | Number of policy exposed to lapse | Number of expected lapse ($B_i$) | A/E lapse ratio | $C_i$ |
|---|---|---|---|---|---|
| C1 | 12.012,00 | 272.308 | 9.620,54 | 1,25 | 280,90 |
| C2 | 1.066,00 | 24.687 | 1.170,55 | 0,91 | 36,85 |
| C3 | 3.874,00 | 88.959 | 3.331,02 | 1,16 | 94,32 |
| C4 | 1.230,00 | 36.998 | 890,00 | 1,38 | 121,04 |
| C5 | 1.803,00 | 44.993 | 1.902,12 | 0,95 | 57,33 |
| C6 | 4,00 | 77 | 2,33 | 1,72 | 0,09 |
| Total | 19.989 | | 16.916,56 | 1,18 | 590,53 |

Using the formula for the variance we obtain that $\hat{\sigma}^2 = 0,03$. Now $Z$ can be calculated according to the formula (7).

*Table 9. Lapse Results by Policy*

| | Overall A/E Ratio | Z | Company A/E Ratio | Number of Lapses | Buhlmann estimates A/E Ratio |
|---|---|---|---|---|---|
| C1 | 118,2% | 0,995 | 124,9% | 12.012 | 124,8% |
| C2 | 118,2% | 0,963 | 91,1% | 1.066 | 92,1% |
| C3 | 118,2% | 0,987 | 116,3% | 3.874 | 116,3% |
| C4 | 118,2% | 0,958 | 138,2% | 1.230 | 137,4% |
| C5 | 118,2% | 0,977 | 94,8% | 1.803 | 95,3% |
| C6 | 118,2% | 0,050 | 171,7% | 4 | 120,8% |

Both approaches assume that the mean value (total A / E ratio) is constant over the time. This assumption for a period of 5 years for mortality can be representative. However, the rates of lapse vary with factors such as economic conditions, as well as events in the field of insurance, including the introduction of new products or changes in regulations. Therefore, many actuaries limit this assumption for a shorter period.

The scores of the models are very similar and vary by less than 5%, as a result of the application of variance.

## 3.1. Summary of greatest accuracy credibility

Bayesian methods express their advantage by using information of all parameters in estimating each individual parameter. In addition, the Bayes method considers all unknown parameters that appear in the statistical model as random variables, and their distribution is linked to the available information.

All aspects of the process, approximation or assumption are clearly defined and stated, after which the solution is derived from the basic principles of probability. The method has a clear objective function, minimizing mean square error. Also, there are no arbitrary choices that are not related to the observed random variables.

However, linear Bayes approximation can give poor results, especially in situations where the random variable has a heavy tail. The second variance, or empirical Bayes estimates $\alpha$ , can be negative, since $\alpha$ is the variance and the real value cannot be negative. A case where a prior distribution is known from some other source represents an exception, thus resulting in a possibility to determine credibility only if there are multiple entity data.

## 4. FUTHER RESEARCH

The credibility theory relies heavily on actuarial, or other related expert judgment, when making decisions in the selection, development and use of credibility. The theory of fuzzy system allows the use of subjective judgments expressed by using vague terms, relations and statements to describe the problem and selection of alternatives in a decision making process. The system based on fuzzy rules leads to the modelling of human interpretation.[467] The nature of human behavior is to judge on the basis of evidence, as a starting point for making adequate decisions and achieving goals.

However, uncertainty and vagueness are the most common reasons for the occurrence of errors in assessing the characteristics and size of certain phenomena, due to the lack of clear and accurate information about the environment. Reduction of the level of assessment subjectivity is thus greatly affected, which is stressed as one of the more significant problems in experience and expert assessment.[468]

Today, there is a common view that when a complex system does not provide enough information to be described wih mathematical expressions that could be implemented within a precise mathematical model, the theory of fuzzy mathematics provides acceptable characteristics compared to other approaches for solving those problems.

Numerous authors in determining premium rate introduce factors of uncertainty, vagueness, or ambiguity. Within the same homogeneous group there may be differences between the influences of the identified risk elements that can individually or collectively affect the overall risk of specific insurance cover. Also, the risks that are accepted in insurance cover are often very complex and there are different causal relationships between known and expected risks within a homogeneous group of insurance. This is often the result of the uncertainty of the evaluators and the variability of the conditions.[469] This fact

[467] Kerkez, M., Ralević, N., Milutinović, O., Vojinović, Ž., & Mladenović-Vojinović, B. (2018). Integrated Fuzzy System and Multi-Expression Programming Techniques for Supplier Selection. *Technical Gazette, 26(1*), p. 1.

[468] Gajović, V., Kerkez, M., & Kočović, J. (2017). Modeling and simulation of logistic processes: risk assessment with a fuzzy logic technique. *Simulation. Transactions of the Society for Modeling and Simulation International*, p. 10.

[469] Tomašević, M., Ralević, N., Stević, Ž., Marković, V., & Tešić, Z. (2018). Adaptive fuzzy model for determination of quality assessment services in supply chain, *Tehnički Vjesnik, 25*(6), (in press).

determines the need to define the gross premium and gross premium rates in a certain range, from minimum to maximum value. In the future research authors will attempt to study the risk using credibility theory focusing on risk situations described by fuzzy variables.