# Bayesian estimates of health care costs in Slovakian insurance companies

P. Jindrová, K. Seinerová

***Abstract*—** This paper presents an example of application of the Bayesian analysis to estimate health care expenses by health insurance companies. In particular, the Bühlmann-Straub model is applied onto real data from three insurance companies in Slovakia in order to calculate the credible health care costs per person and the corresponding credibility factors for each company. The results show that while the largest insurance company can reliably predict the health care costs for the subsequent time period based on its own statistical data, the smaller companies are unable obtain meaningful predictions without involving the data from the whole marked. The presented case study shows that the Bayesian analysis is a useful tool for calculations in the field of health insurance.

***Keywords*—**Credibility Model, Bayesian Estimation, Health Insurance**.**

## I. Introduction

NOWADAYS, questions related to funding of the health care in post-communist countries are discussed. All residents of Slovakia, including foreigners living or being employed in Slovakia, have to pay insurance premium as a contribution to the public health insurance. Thanks to health insurance system there is availability of health care for all the insured persons, regardless of the height of their payments.

There are three health insurance companies in Slovakia - Všeobecná zdravotná poisťovna, Dôvera zdravotná poisťovňa, and Union poisťovňa.

Figure 1 presents the number of insured persons in the time period 2010-2014 for each Slovakian health insurance company. Evidently, Všeobecná zdravotná poisťovňa insures major part of Slovakian population, the number of persons it insures is even higher than sum of persons insured by both other health insurance companies. During the considered time period the number of persons insured by Všeobecná zdravotná poisťovna decreases moderately, especially between 2011 and 2012.

To be able to ensure financial means, every health insurance company needs to appraise health care expenses for the subsequent time period. It is appropriate to use information about health care expenses based on the data from all the health care insurance companies in the country. For this purpose, it is necessary to estimate these expenses by applying mathematical and probabilistic models, for example those based on Bayesian analysis. Other possibilities of using Bayesian estimates in insurance are discussed in [2]-[6].

The time series of the total health care expenses and number of insured persons are the basis for estimates the expenses of the health insurance companies in Slovakia. Data from all three Slovakian health insurance companies covering time period 2010-2014 are used for computation in this paper. Bühlmann-Straub model is applied to estimate the health care costs per person for the subsequent time period. Estimation of the health care costs for each health insurance company is a linear combination of estimated costs based on inner historical data of this insurance company and estimations based on outer information related to health care costs from all insurance companies. This approach, taking both own and the undertaken information into account, guaranties objective estimates of the health care costs.

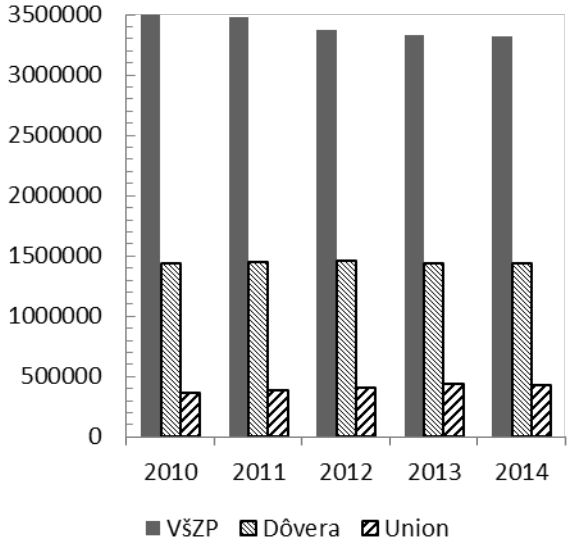


Fig. 1 Number of insurered persons

This work was supported by the project SGSFES_2015001 Ekonomický a sociální rozvoj v soukromém a veřejném sektoru (Economical and social development in private and public sector) financed from Czech Republic funds.

P. Jindrová is with Institute of Mathematics and Quantitative Methods, Faculty of Economics and Administration, University of Pardubice, Pardubice, Studentská 84, 532 10 Pardubice, Czech Republic (e-mail: Pavla.Jindrova@upce.cz).

K. Seinerová is with Institute of Mathematics and Quantitative Methods, Faculty of Economics and Administration, University of Pardubice, Pardubice, Studentská 84, 532 10 Pardubice, Czech Republic (e-mail: Katerina@Seinerova@upce.cz).

## II. DATA AND METHODOLOGY

The health care insurance in Slovakia is provided by three health insurance companies. Computations presented in this paper are based on data extracted from their yearbooks covering time period 2010-2014, available on websites [9]-[11].

Bühlmann-Straub model used in this paper follows here.

Let the insurance company, for which the estimates of health care expenses (resp. net insurance premium) are computed, is one of $N$ insurance companies. Total expenses in previous $n$ years are known for all these insurance companies.

Let $Y_{ij}$ is a variable, describing total health care costs of the $i$-th health insurance company in Slovakia ($i$ = 1, …, $N$ = 3) in the $j$-th year ($j$ = 1, …, $n$ = 5).

Let $P_{ij}$ stands for the number of insured person for each health insurance company $i$ = 1, …, 3 and each year $j$ = 1, …, 5.

According to [7], standardized health care costs $X_{ij} = Y_{ij} / P_{ij}$, satisfy the following conditions:

- For every $i$ = 1, …, $N$ the distribution of variable $X_{ij}$ depends on an unknown parameter $\theta_i$, equal for all the years $j$ = 1, …, $n$.
- For every $j$ = 1, …, $n$ the variables $X_{i1} / \theta_i$, $X_{i2} / \theta_i, ..., X_{in} / \theta_i$ are independent, but not necessarily identically distributed.

Then, two functions depending on $j$ can be defined as:

$$E\left(X_{ij} / \theta_i\right) = m\left(\theta_i\right) \tag{1}$$

$$D\left(X_{ij} / \theta_i\right) = \frac{s^2\left(\theta_i\right)}{P_{ij}} \tag{2}$$

The above mentioned relations express conditions satisfied for every risk $i$ = 1, …, $N$. Relationships between risks are described by the following conditions:

- Parameters of risks $\theta_1,...,\theta_N$ are random variables, which are independent and identically distributed.
- For $i \neq k$ are $\left(\theta_i, X_{ij}\right)$ and $\left(\theta_k, X_{km}\right)$ independent.

Because parameters of risk $\theta_1,...,\theta_N$ are identically distributed, values $E\left(m\left(\theta_i\right)\right),\ D\left(m\left(\theta_i\right)\right),\ E\left(s^2\left(\theta_i\right)\right)$ are independent on $i$ and can be denote as $E\left(m\left(\theta\right)\right),\ D\left(m\left(\theta\right)\right),\ E\left(s^2\left(\theta\right)\right)$.

Following formulas are used to simplify the calculation:

$$P_i = \sum_{j=1}^{n} P_{ij} \tag{3}$$

$$P = \sum_{i=1}^{N} P_i \tag{4}$$

$$Y_i = \sum_{j=1}^{n} Y_{ij} \tag{5}$$

$$Y = \sum_{i=1}^{N} Y_i \tag{6}$$

$$\overline{X}_i = \frac{1}{P_i} \sum_{j=1}^{n} P_{ij} X_{ij} = \frac{1}{P_i} \sum_{i=1}^{n} Y_{ij} \tag{7}$$

$$\overline{X} = \frac{1}{P} \sum_{i=1}^{N} \sum_{j=1}^{n} P_{ij} X_{ij} = \frac{1}{P} \sum_{i=1}^{N} P_i \overline{X}_i \tag{8}$$

$$P^* = \frac{1}{Nn-1} \sum_{i=1}^{N} P_i \cdot \left(1 - \frac{P_i}{P}\right) \tag{9}$$

Then according to [1], the rules for estimating the parameters $E\left(m\left(\theta\right)\right), D\left(m\left(\theta\right)\right), E\left(s^2\left(\theta\right)\right)$ are:

$$estE(m(\theta)) = \overline{X} \tag{10}$$

$$estE(s^2(\theta)) = \frac{1}{N(n-1)} \sum_{i=1}^{N} \sum_{j=1}^{n} P_{ij} \left(X_{ij} - \overline{X}_i\right)^2 \tag{11}$$

$$estD(m(\theta)) = \frac{1}{P^*} \left\{ \begin{array}{l} \frac{1}{Nn-1} \sum_{i=1}^{N} \sum_{j=1}^{n} P_{ij} \left(X_{ij} - \overline{X}\right)^2 \\ -\frac{1}{N(n-1)} \sum_{i=1}^{N} \sum_{j=1}^{n} P_{ij} \left(X_{ij} - \overline{X}_i\right)^2 \end{array} \right\} \tag{12}$$

Credibility factor for the $i$-th risk is calculated in form

$$Z_i = \frac{P_i}{P_i + \frac{E(s^2(\theta))}{D(m(\theta))}} \tag{13}$$

The value of credibility factor shows the rate of reliability of own data for every health insurance company.

Estimates of parameters $E\left(m\left(\theta\right)\right), D\left(m\left(\theta\right)\right), E\left(s^2\left(\theta\right)\right)$ are the same for all the insurance companies. Credibility factor $Z_i$ differs for each company. The higher is the value of risk rate $P_i$, the higher is the value of credibility factor $Z_i$.

Then, according to [8], for the estimation of credible health care for the $i$-th risk, the following formula is used.

$$\begin{aligned} E(m(\theta) / X) &= Z_i \overline{X}_i + (1 - Z_i) E(m(\theta)) \\ &= Z_i \overline{X}_i + (1 - Z_i) \overline{X} \end{aligned} \tag{14}$$

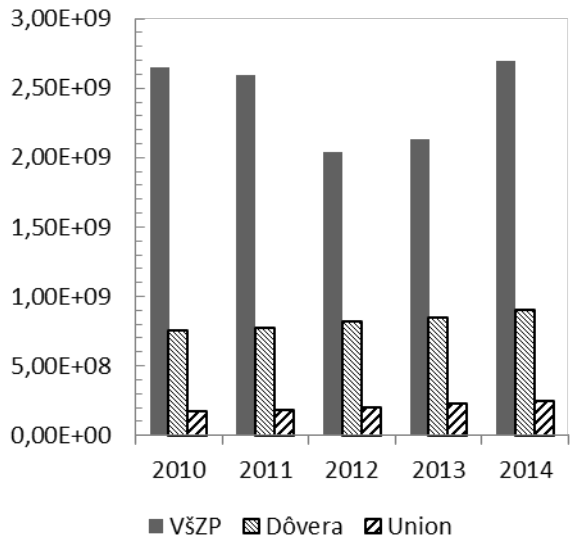


Fig. 2 Health care costs

## III. Results

Bühlmann-Straub model was applied to compute credible net premium, here used as health care cost per person. The computation was based on a set of real data published in yearbooks of Slovakian health insurance companies.

The basis for the computation of credible health care costs per person according to formula (14) was the data from all Slovakian health insurance companies in years 2010-2014, namely total health care costs paid by particular health insurance company ($Y_{ij}$) and amount of persons insured by particular health insurance company ($P_{ij}$). These values are shown in Fig. 1 and Fig. 2. Amount of insured persons in each health insurance company and each year presents the rate of insurance risk.

Total amount of insured persons $P_i$ insured by each insurance company $i$ in the whole period, total costs $Y_i$ paid by each insurance company in the whole period, and average insurance costs per person in the whole period $\overline{X}_i$ for each insurance company were computed according to (3), (5) and (7).

Results of computation of these characteristics are presented in Table I.

Table I Chosen characteristics

| | $P_i$ | $Y_i$ | $\overline{X}_i$ |
|---|---|---|---|
| VšZP | 16 995 530 | 12 111 980 900 | 712.66 |
| Dôvera | 7 220 865 | 4 093 000 000 | 566.83 |
| Union | 2 034 385 | 1 029 760 630 | 506.18 |

Source: own calculation

Then, following values are computed:

$$P = \sum_{i=1}^{3} P_i = 26250780$$

$$Y = \sum_{i=1}^{3} Y_i = 17\,234\,741530$$

$$\overline{X} = 656.54$$

Then values of $X_{ij} = \frac{Y_{ij}}{P_{ij}}$, representing standardized health care costs per person in each year and each insurance company (in €). These values are then used to estimate parameters of the model according to (10)-(12).

$$estE\left(m(\theta)\right) = 656.54$$
$$estE\left(s^2(\theta)\right) = 9318644226$$
$$estD\left(m(\theta)\right) = 10607.25$$

According to (13), values of credibility factors $Z_i$ are computed for every health insurance company. It is necessary to find these values to be able to compute credible estimations of net insurance costs per person according to (14). Values of credibility factors and credible health care costs are presented in Table II. The value of credibility factor $Z_i$ shows the effect of inner data from the particular insurance company on the value of the credible health care costs per person. (1-$Z_i$) shows the same effect for the outer data from all the insurance companies.

Values of credible estimates of the health care costs per person can be used for appraising net insurance premium for the following year in every insurance company. Since the data used for computations covers period 2010-2014, computed values of credible estimates of average costs are estimates of net health care costs per person for each insurance company in 2015.

Table II Credibility factors $Z_i$ and credible health care costs per person (in €)

| | $Z_i$ | Credible health care costs |
|---|---|---|
| VšZP | 0.9508 | 709.90 |
| Dôvera | 0.8915 | 576.56 |
| Union | 0.6984 | 551.53 |

Source: own calculation

Table II shows the differences in values of credibility factors $Z_i$ among the Slovakian health insurance companies, and subsequently big differences in estimates of net insurance costs per person in Euros. Value $Z_i$ = 0.9508 for major Slovakian insurer VšZP shows, that this health insurance company can rely from 95.08 % on its own. The second major insurer Dôvera has the credibility factor 0.8915. This insurance company can rely from 89.15 % on its own data and from 10.85 % on the information from all the insurance companies. Even greater effect of outer data on credible health

care costs is evident at the smallest health insurer Union. Health insurance company Union have to rely on outer data from more than 30 %.

Estimated values of net insurance costs show big differences between the three insurance companies. The origin of these differences falls beyond the scope of this paper.

## IV. Conclusion

Every health insurance company needs to appraise its costs for following time period. Estimates of these costs can be computed by applying Bayesian analysis, especially Bayesian models of credibility. Results obtained by applying Bühlmann-Straub model on real data published by Slovakian health insurance companies, namely height of health insurance costs paid by insurance companies and amount of persons insured by these companies in years 2010-2014, show differences in estimations of health care costs per person among the insurance companies. Applying Bühlmann-Straub model on considered data, values of credibility factor are computed for each Slovakian health insurance company. These values show the effect of the inner data from particular insurance companies on the value of credible health care costs per person. All insurance companies should bother both inner and outer data, however, the majority effect on the results relates to the inner data of particular health insurance company.

The computed values of the regional credible health care costs can be used to estimate the real health care costs for the following time period in every insurance company. This information is very important for estimating financial means. Using the Bayesian analysis seems to be useful for computations in the field of the health insurance.

## References

[1] P. J. Boland, *Statistical and Probabilistic Methods in Actuarial Science*. New York: Taylor & Francis Group. 2007.

[2] J. Gogola, Spôsob permanentnej úpravy výšky pojistného v neživotnom poistení. *E + M Economics and Management*, vol. 16 (4), 2013, pp. 134-141.

[3] P. Jindrová, Credibility Risk Models in Accident Insurance. In: *Proceedings of the 7th International Scientific Conference: Managing and Modelling of Financial risk*. Ostrava: VŠB – TU Ostrava, Faculty of Economics, Finance department. 2014, pp. 307-316.

[4] P. Jindrová, V. Pacáková, Bayesian Probability Models for Critical Illness Insurance. In *Recent Advances in Mathematical Methods in Applied Sciences*. Athens: WSEAS Press, 2014. pp. 218-221.

[5] P. Jindrová, V. Pacáková, Actuarial Models for Valuation of Critical Illness Insurance. *International Journal of Mathematical Models and Methods in Applied Sciences*, vol. 9, 2015, pp. 218-226.

[6] P. Jindrová,,K. Seinerová, Bayesian Estimates of the Regional Costs in Public Health System of the Czech Republic. In: *Proceedings of the 12th International Scientific Conference: European Financial Systems 2015*, Brno: Masaryk University, 2015.

[7] V. Pacáková, *Aplikovaná poistná štatistika*, 3th ed. Bratislava: Iura Edition. 2004.

[8] V. Pacáková, Bayesian Estimations in Insurance Theory and Practice. In *Advances in Mathematical and Computational Methods: proceedings of the 14th WSEAS International Conference on Mathematical and Computational methods in Science and Engeneering*. Stevens Point: WSEAS Press, 2012. pp. 127-131.

[9] Správy o hospodárení – Dôvera zdravotná poisťovňa (2015). Available: http://www.dovera.sk/sprava-o-hospodareni

[10] Výroční správy – Union zdravotná poisťovňa, a.s. (2015). Available: http://www.union.sk/union-zp-vyrocne-spravy

[11] Výroční správy – Všeobecná zdravotná poisťovňa (2015). Available: http://www.vszp.sk/o-nas/vyrocna-sprava/

**Mgr. P. Jindrová**, **Ph.D.** graduated from Mathematical analysis at the Faculty of Science of Palacky University in Olomouc in 1993. She lectures at Institute of Mathematics and Quantitative Methods at Faculty of Economic and Administration in Pardubice since 2004. She finished her dissertation thesis which deals with problems of aggregation and disaggregation in economics and mathematics in 2011 at the same faculty.