Optimization approaches to multiplicative tariff of rates estimation in non-life insurance

Martin Branda

Charles University in Prague
Faculty of Mathematics and Physics
Department of Probability and Mathematical Statistics
Sokolovská 83, 186 00 Prague, Czech Republic
tel.: +420 221 913 343, fax: +420 222 323 316, e-mail: branda@karlin.mff.cuni.cz
&
Kooperativa pojišťovna, a.s., Vienna Insurance Group
Head Office, Actuarial Department
Pobřežní 665/21, 186 00 Prague, Czech Republic

Abstract

We focus on rating of non-life insurance contracts. We employ multiplicative models with basic premium levels and specific surcharge coefficients for various levels of selected risk/rating factors. We use generalized linear models (GLM) to describe the probability distribution of total losses for a contract during one year. We show that the traditional frequency–severity approaches based only on GLM with logarithmic link function can lead to estimates which do not fulfill business requirements. For example, a maximal surcharge and monotonicity of coefficient can be desirable. Moreover, our approach can handle total losses, which are based on arbitrary loss distributions, possibly decomposed into several classes, e.g. small and large or property and bodily injury. Various costs and loadings can be also incorporated into the tariff rates. We propose optimization problems for rate estimation which enable to hedge against expected losses and to take into account a prescribed loss ratio and other business requirements. Moreover, we introduce a stochastic programming problem with reliability type constraints which incorporate riskiness of each rate cell. In the numerical study, we apply the approaches to Motor Third Party Liability (MTPL) policies.

Keywords: non-life insurance, rate making, generalized linear models, optimization model, stochastic programming, MTPL.

1. Introduction

Estimation of prices for which policies are sold is a highly important task for insurance companies. In this paper, we will focus on rating of non-life insurance contracts. Traditional credibility models take into account known history of a policyholder and project it into policy rate, see, e.g., Bühlmann and Gisler (2005). However, for new business, i.e. clients coming for a new insurance policy, the history need not to be known or the information may not be reliable. Thus traditional approaches of credibility theory can not be used. We will employ models which are based on settled claims of new contracts from the previous years. This experience is transferred using generalized linear models (GLM), see McCullagh and Nelder (1989), which cover many important regression models used in insurance, cf. Antonio and Beirlant (2007), Denuit et al. (2007), de Jong and Heller (2008), Ohlsson (2008), Ohlsson, and Johansson (2010). The frequency–severity approach is the most frequently used, where expected claim count on a policy during one year and expected claim size can be explained by various independent variables, which can serve as segmentation criteria, e.g. age and gender of the policyholder and region where he or she lives, properties of the object. Using these criteria and GLM with the logarithmic link function we can derive directly basic premium levels and surcharges which enable to take into account riskiness of each policyholder. However, as we will show in this paper, these coefficients need not to fulfill business requirements, for example restriction on maximal surcharge. Moreover, if other link functions are used or the regression dependence is more difficult, optimization models must be employed to set the basic premium levels and the surcharge coefficients.

Stochastic programming techniques can be used to solve optimization problems where random vectors appear. They have already found several applications in insurance, see, e.g., Ermoliev et al. (2000), Hilli et al. (2011). In this paper, we will employ a formulation based on reliability type constraints such as chance constraints and the reformulation based on the one-sided Chebyshev's inequality. The distribution of the random parts will be represented by compound Gamma–Poisson and Inverse Gaussian–Poisson distributions with parameter estimates based on generalized linear models. It can be shown that the Chebyshev's inequality produces bound, which is tight with respect to the distributions with the given expectation and variance, see Chen et al. (2011). Preliminary results on this topic were presented by Branda (2012d) where simple models were presented and the logarithmic transformation used also in this paper was suggested.

The main advantages of our optimization approach can be summarized in the following points:

- GLM with other than logarithmic link function can be used,
- business requirements on surcharge coefficients can be ensured,
- total losses can be decomposed and modeled using different models, e.g. for standard and large losses or for bodily injury and property damage,
- other modelling techniques than GLM can be used to estimate the distribution of total losses over one year, e.g. generalized additive models, classification and regression trees,
- costs and loadings (commissions, tax, office expenses, unanticipated losses, cost of reinsurance) can be incorporated when our goal is to optimize the combined ratio instead of the loss ratio, we obtain final office premium as the output,
- not only the expectation of total losses can be taken into account but also the shape of the distribution, i.e. riskiness can be projected into the final rates,
- the ruin probability can be controlled for the whole portfolio.

This paper is organized as follows. We will propose basic notation in Section 2. In Section 3, we will review definition and basic properties of generalized linear models. We will recall a rate-making approach based directly on GLM. In Section 4, optimization models for rates estimation are introduced which enable to take into account various business requirements and the other generalizations proposed above. These models are extended using stochastic programming techniques in Section 5. Section 6 concludes the paper with an application of the proposed methods to MTPL contracts.

2. Notation and preliminaries

Policies that belong to the same class for each rating factor are said to belong to the same tariff cell and are given the same premium. We denote by $i_0 \in \mathcal{I}_0$ the levels of basic segmentation criterion, e.g. tariff cells, and by $i_1 \in \mathcal{I}_1, \ldots, i_S \in \mathcal{I}_S$ the levels of other segmentation criteria which should help us to take into account underwriting risk, which can be significantly different for each class. We will denote one risk cell $I = (i_0, i_1, \ldots, i_S)$ with

$I \in \mathcal{I} = \mathcal{I}_0 \otimes \mathcal{I}_1 \otimes \cdots \otimes \mathcal{I}_S$. Let W_I denote the number of contracts in the rate cell I. Let aggregated losses over one year for risk cell I be

$$L_I^T = \sum_{w=1}^{W_I} L_{I,w},\ L_{I,w} = \sum_{n=1}^{N_{I,w}} X_{I,n,w},$$

where $N_{I,w}$ is the random number of claims for a contract during one year and $X_{I,n,w}$ is the random claim severity. All the considered random variable are assumed to be independent. For each I, we assume that the random variables $N_{I,w}$ has the same distribution for all w, and $X_{I,n,w}$ for all n and w. We denote by N_I, X_I independent copies of $N_{I,w}, X_{I,n,w}$. Then, the following well-known formulas can be obtain for the mean and the variance of the aggregated losses:

$$\begin{aligned}
\mu_I &= \mathbb{E}[L_I] = \mathbb{E}[N_I]\mathbb{E}[X_I], \\
\mu_I^T &= \mathbb{E}[L_I^T] = W_I \mu_I, \\
\sigma_I^2 &= var(L_I) = \mathbb{E}[N_I]var(X_I) + (\mathbb{E}[X_I])^2 var(N_I), \\
(\sigma_I^T)^2 &= var(L_I^T) = W_I \sigma_I^2.
\end{aligned}$$

We denote the total premium $TP_I = W_I Pr_I$ for the risk cell I. We assume that the risk (office) premium is composed in a multiplicative way from basic premium levels Pr_{i_0} and nonnegative surcharge coefficients $e_{i_1}, \ldots, e_{i_S}$, i.e. we obtain the decomposition

$$Pr_I = Pr_{i_0} \cdot (1 + e_{i_1}) \cdot \cdots \cdot (1 + e_{i_S}).$$

Our goal is to find optimal basic premium levels and surcharge coefficients with respect to a prescribed loss ratio $\hat{LR}$, i.e. to fulfill the random constraints

$$\frac{L_I^T}{TP_I} \leq \hat{LR} \text{ for all } I \in \mathcal{I}. \tag{1}$$

The goal loss ratio $\hat{LR}$ is usually based on a management decision. It is possible to prescribe different loss ratios for each tariff cell but this is not considered in this paper. Note that the relation (1) is influenced by the exposure of the risk cell W_I, since the total losses are considered as a random variable. We can compute mean and variance of the ratio

$$\begin{aligned}
\mathbb{E}\left[\frac{L_I^T}{TP_I}\right] &= \frac{\mathbb{E}[L_I^T]}{W_I Pr_I} = \frac{\mathbb{E}[N_I]\mathbb{E}[X_I]}{Pr_I}, \\
var\left(\frac{L_I^T}{TP_I}\right) &= \frac{var(L_I^T)}{W_I^2 Pr_I^2} = \frac{\mathbb{E}[N_I]var(X_I) + (\mathbb{E}[X_I])^2 var(N_I)}{W_I Pr_I^2}.
\end{aligned}$$

Usually, the expected value of the loss ratio is bounded

$$\frac{\mathbb{E}[L_I^T]}{TP_I} = \frac{\mathbb{E}[L_I]}{Pr_I} \leq \hat{LR} \text{ for all } I \in \mathcal{I}. \tag{2}$$

If $\hat{LR} = 1$, we obtain the netto-premium. However, this approach does not take into account riskiness of each tariff cell. A natural requirement can be be that the inequalities (1) are fulfilled with a prescribed probability leading to separate chance (probabilistic) constraints

$$P\left(\frac{L_I^T}{TP_I} \leq \hat{LR}\right) \geq 1 - \varepsilon, \text{ for all } I \in \mathcal{I}, \tag{3}$$

where $\varepsilon \in (0, 1)$, usually ε is small. Another probabilistic approach bounds the loss ratio over the whole line of business:

$$P\left(\frac{\sum_{I\in\mathcal{I}} L_I^T}{\sum_{I\in\mathcal{I}} TP_I} \leq \hat{LR}\right) \geq 1 - \varepsilon.$$

How the risk is allocated to tariff cells will be discussed later in this paper.

3. Rate-making using generalized linear models

In this section, we introduce generalized linear models (GLM), which cover many regression models useful in insurance. GLM are based on the following three building blocks:

1) The dependent variable Y_i has a distribution from the exponential family with the probability density function

$$f(y; \theta_i, \varphi) = \exp\left\{\frac{y\theta_i - b(\theta_i)}{\varphi} + c(y, \varphi)\right\},$$

where b, c are known functions and θ_i, φ are unknown canonical (dependent on observation) and dispersion (common for all observations) parameters.

2) A linear combination of independent variables is considered

$$\eta_i = \sum_j X_{ij}\beta_j,$$

where β_j are unknown parameters and X_{ij} are given values of predictors.

3) The dependency is described by a link function g which is strictly monotonous and twice differentiable

$$\mathbb{E}[Y_i] = \mu_i = g^{-1}(\eta_i).$$

The most important members of the exponential family are proposed in Table 1 including basic characteristics, which are introduced below. The following relations can be obtained for the expectation and variance under the assumption that b is twice continuously differentiable

$$\begin{aligned} \mathbb{E}[Y] &= b'(\theta), \\ var(Y) &= \varphi b''(\theta) = \varphi V(\mu), \end{aligned}$$

where the last expression is rewritten using the variance function which is defined as $V(\mu) = b''[(b')^{-1}(\mu)]$, i.e. the variance depends on the mean only.

Distribution	Density $f(y;\theta,\varphi)$	Dispersion param. φ	Canonical param. $\theta(\mu)$	Mean value $\mu(\theta)$	Variance function $V(\mu)$
$Po(\mu)$	$\frac{\mu^y e^{-\mu}}{y!}$	1	$\ln(\mu)$	e^θ	μ
$\Gamma(\mu,\nu)$	$\frac{1}{\Gamma(\nu)y}\left(\frac{y\nu}{\mu}\right)^\nu e^{-\frac{y\nu}{\mu}}$	$\frac{1}{\nu}$	$-\frac{1}{\mu}$	$-\frac{1}{\theta}$	μ^2
$IG(\mu,\lambda)$	$\sqrt{\frac{\lambda}{2\pi y^3}}e^{-\frac{\lambda(y-\mu)^2}{2\mu^2 y}}$	$\frac{1}{\lambda}$	$-\frac{1}{2\mu^2}$	$\frac{1}{\sqrt{-2\theta}}$	μ^3

Table 1: Examples of distributions from the exponential family

Maximum likelihood method is used to estimate the parameters of GLM. Overdispersion is a phenomenon that is often observed in practice where the variance need not to be equal to the expected value as it is for the Poisson distribution. In this case, the dispersion parameter φ is not set to 1 but is estimated from data. The packages for GLM estimation usually offer an overdispersed Poisson model or a negative-binomial model. Quasi-likelihood function must be used for the overdispersed Poisson model, see McCullagh and Nelder (1989). Another interesting class of models are zero-inflated models, see Cameron and Trivedi (1998).

3.1. Pure premium estimation

Although the losses L_I are random, the simplest way, which is often used in practice, is to hedge against the expected value of aggregated losses (2).

This can be done directly using GLM with the logarithmic link function $g(\mu) = \ln \mu$. Poisson and Gamma or Inverse Gaussian regressions without an intercept can be used to estimate the parameters for the expected number of claims and claims severity. If we use the logarithmic link function in both regression models, then we can get for each $I = (i_0, i_1, \dots, i_S)$

$$\begin{aligned}
\mathbb{E}[N_I] &= \exp\{\lambda_{i_0} + \lambda_{i_1} + \cdots + \lambda_{i_S}\}, \\
\mathbb{E}[X_I] &= \exp\{\gamma_{i_0} + \gamma_{i_1} + \cdots + \gamma_{i_S}\},
\end{aligned}$$

where λ_i, γ_i are the estimated coefficients. Thus for the expected loss it holds

$$\mathbb{E}[L_I] = \exp\{\lambda_{i_0} + \gamma_{i_0} + \lambda_{i_1} + \gamma_{i_1} + \cdots + \lambda_{i_S} + \gamma_{i_S}\}.$$

The basic premium levels and the surcharge coefficient are based on a product of normalized coefficients. They can be estimated as

$$\begin{aligned}
Pr_{i_0} &= \frac{\exp\{\lambda_{i_0} + \gamma_{i_0}\}}{\hat{LR}} \cdot \prod_{s=1}^{S} \min_{i \in \mathcal{I}_s} \exp(\lambda_i) \cdot \prod_{s=1}^{S} \min_{i \in \mathcal{I}_s} \exp(\gamma_i), \\
e_{i_s} &= \frac{\exp(\lambda_{i_s})}{\min_{i_s \in I_s} \exp(\lambda_{i_s})} \cdot \frac{\exp(\gamma_{i_s})}{\min_{i_s \in I_s} \exp(\gamma_{i_s})} - 1,
\end{aligned}$$

Under this choice, the constraints (2) are fulfilled with respect to the expectations. Note that if the less risky classes are selected as the reference categories, the normalization above is not necessary. If the models are estimated using historical data, it is important to incorporate inflation of the losses. In our case, it is possible to inflate the basic premium levels only.

The approach above is highly dependent on using GLM with the logarithmic link function. It can be hardly used if other link functions are used, interaction or other regressors than the segmentation criteria are considered. For the aggregated losses modelling, we can employ models with the logarithmic link and with a Tweedie distribution for $1 < p < 2$, which correspond directly to the compound Poisson–gamma distributions. The expected loss is explained by the rating factors only.

However, the surcharge coefficient estimated by both methods above often violate business requirements, especially they can be too high, as we will show in the numerical study. Then optimization models can be the only way how to obtain the basic premium levels as well as the surcharge coefficient. The loss modeling can be split by claim type, e.g. different models can be used for standard and large losses or for bodily injury and property damage. Moreover, the riskiness can be taken into account as we will show in Section 5.

4. Optimization problem for rate estimation

Starting from this section, we can assume that L_I contains not only losses but also various costs and loadings, thus we can construct the tariff rates with respect to a prescribed combined ratio. For example, the total loss can be composed as follows

$$L_I = (1 + vc_I)\big[(1 + inf_s)L_I^s + (1 + inf_l)L_I^l\big] + fc_I,$$

where small L_I^s and large claims L_I^l are modeled separately, inflation of small claims inf_s and large claims inf_l, proportional costs vc_I and fixed costs fc_I are incorporated into total losses.

The constraints (2) with expectation can be rewritten as

$$\mathbb{E}[L_{i_0,i_1,\dots,i_S}] \le \hat{LR} \cdot Pr_{i_0} \cdot (1 + e_{i_1}) \cdot \dots \cdot (1 + e_{i_S}). \tag{4}$$

There can be prescribed a business limitation that the highest aggregated risk surcharge is lower than a given level $r^{max} \ge 0$. It is also possible to set an upper bound on each surcharge coefficient. We would like to minimize basic premium levels and surcharges, which are necessary to fulfill the prescribed loss ratio and the business requirements. The premium is minimized to ensure maximal competitiveness on a market. This can be further strengthened by discounts, which are not in the scope of this paper. We obtain the following nonlinear optimization problem where the premium is minimized under the condition that the premium covers the expected losses with respect to the prescribed loss ratio and that the maximal possible surcharge is less than the prescribed level r^{max}:

$$\begin{aligned} \min \sum_{I \in \mathcal{I}} w_I Pr_{i_0} (1 + e_{i_1}) \cdot \dots \cdot (1 + e_{i_S}) & \\ \text{s.t.} & \\ \hat{LR} \cdot Pr_{i_0} \cdot (1 + e_{i_1}) \cdot \dots \cdot (1 + e_{i_S}) &\ge \mathbb{E}[L_{i_0,i_1,\dots,i_S}], \\ (1 + e_{i_1}) \cdot \dots \cdot (1 + e_{i_S}) &\le 1 + r^{max}, \\ e_{i_1}, \dots, e_{i_S} &\ge 0,\ (i_0, i_1, \dots, i_S) \in \mathcal{I}. \end{aligned} \tag{5}$$

This problem is nonlinear nonconvex, thus very difficult to solve. However, using the logarithmic transformation of the decision variables $u_{i_0} = \ln(Pr_{i_0})$ and $u_{i_s} = \ln(1 + e_{i_s})$ and by setting

$$b_{i_0,i_1,\dots,i_S} = \ln(\mathbb{E}[L_{i_0,i_1,\dots,i_S}]/\hat{LR}),$$

the problem can be rewritten as a nonlinear convex programming problem, which can be efficiently solved by standard software tools:

$$\begin{aligned} \min \sum_{I \in \mathcal{I}} w_I e^{u_{i_0}+u_{i_1}+\cdots+u_{i_S}} & \\ \text{s.t.} & \\ u_{i_0}+u_{i_1}+\cdots+u_{i_S} &\geq b_{i_0,i_1,\ldots,i_S}, \\ u_{i_1}+\cdots+u_{i_S} &\leq \ln(1+r^{max}), \\ u_{i_1},\ldots,u_{i_S} &\geq 0,\ (i_0,i_1,\ldots,i_S) \in \mathcal{I}. \end{aligned} \tag{6}$$

The problems (5) and (6) are equivalent in the following sense: $\hat{Pr}_{i_0}, \hat{e}_{i_1}, \ldots, \hat{e}_{i_S}$ is an optimal solution of the problem (5) if and only if $\hat{u}_{i_0}, \hat{u}_{i_1}, \ldots, \hat{u}_{i_S}$ is an optimal solution of the problem (6) with the relation $\hat{u}_{i_0} = \ln(\hat{Pr}_{i_0})$ and $\hat{u}_{i_s} = \ln(1+\hat{e}_{i_s})$. Moreover, the estimates do depend on the exposures of the tariff cells.

4.1. Optimization over a net of coefficients

In this section, we will outline how to modify the previous optimization model to the case when the surcharge coefficient are selected from a discrete set of values. For simplicity we assume that the coefficients are selected from an equidistant net. Let $r_s > 0$ be a step, usually 0.1 or 0.05. Then the surcharge coefficient can be modelled as

$$e_{i_s} = x_{i_s} \cdot r_s,$$

where $x_{i_s} \in \{0, \ldots, J_s\}$ are discrete variables and $J_s = \lfloor r^{max}/r_s \rfloor$. However, after logarithmic transform we obtain a problem, which is hardly solvable. Therefore, we will use another formulation using new binary variables. We set

$$u_{i_s} = \sum_{j=0}^{J_s} y_{i_s,j} \ln(1 + j \cdot r_s),$$

together with a condition

$$\sum_{j=0}^{J_s} y_{i_s,j} = 1,$$

which ensures that at least one coefficient value is selected.

5. Stochastic programming problems for rate estimation

In this section, we propose stochastic programming formulations which take into account compound distribution of random losses not only the expected value. We employ chance constraints for satisfying the constraints (1) with prescribed levels. However, chance constrained problems are very computationally demanding in general and various approximation methods are usually employed, see Branda (2012A, 2012B, 2012C), Prékopa, A. (1995, 2003) for various solution approaches and possible reformulations. Reliability constraints were also discussed by Nemirovski and Shapiro (2006).

5.1. Individual risk model

If we prescribe a small probability level $\varepsilon \in (0,1)$ for violating the loss ratio in each tariff cell, we obtain the following chance (probabilistic) constraints

$$P\left(L^T_{i_0,i_1,\ldots,i_S} \le \hat{LR} \cdot W_{i_0,i_1,\ldots,i_S} \cdot Pr_{i_0} \cdot (1+e_{i_1}) \cdot \cdots \cdot (1+e_{i_S})\right) \ge 1-\varepsilon,$$

which can be rewritten using quantile function of $L^T_{i_0,i_1,\ldots,i_S}$ as

$$\hat{LR} \cdot W_{i_0,i_1,\ldots,i_S} \cdot Pr_{i_0} \cdot (1+e_{i_1}) \cdot \cdots \cdot (1+e_{i_S}) \ge F^{-1}_{L^T_{i_0,i_1,\ldots,i_S}}(1-\varepsilon).$$

By setting

$$b_I = \ln\left[\frac{F^{-1}_{L^T_I}(1-\varepsilon)}{W_I \cdot \hat{LR}}\right],$$

the formulation (6) can be used. However, it can be very difficult to compute the quantiles $F^{-1}_{L^T_I}$ for the compound distributions, see, e.g., Withers and Nadarajah (2011), and Central Limit Theorem can not be used, since the exposition can be too low. Instead of approximating the quantiles, we can employ the one-sided Chebyshev's inequality based on the mean and variance of the compound distribution resulting in the constraints

$$P\left(\frac{L^T_I}{TP_I} \ge \hat{LR}\right) \le \frac{1}{1+(\hat{LR}\cdot TP_I - \mu^T_I)^2/(\sigma^T_I)^2} \le \varepsilon, \tag{7}$$

for $\hat{LR} \cdot TP_I \ge \mu^T_I$. Chen et al. (2011) showed that the bound is tight for all distributions $\mathcal{D}$ with the expected value μ^T_I and the variance $(\sigma^T_I)^2$, i.e.

$$\sup_{\mathcal{D}:\ \mathbb{E}[L^T_I]=\mu^T_I,\ var(L^T_I)=(\sigma^T_I)^2} P\big(L^T_I \ge \hat{LR}\cdot TP_I\big) = \frac{1}{1+(\hat{LR}\cdot TP_I - \mu^T_I)^2/(\sigma^T_I)^2},$$

for $\hat{LR} \cdot TP_I \geq \mu_I^T$. Thus, the constraint can be seen as robust with respect to the distribution with a given mean and variance, i.e. it is ensured that

$$\sup_{\mathcal{D}:\ \mathbb{E}[L_I^T]=\mu_I,\ var(L_I^T)=(\sigma_I^T)^2} P\big(L_I^T \geq \hat{LR} \cdot TP_I\big) \leq \varepsilon.$$

The inequality (7) leads to the following constraints, which serve as conservative approximations:

$$\mu_I^T + \sqrt{\frac{1-\varepsilon}{\varepsilon}}\sigma_I^T \leq \hat{LR} \cdot TP_I.$$

Finally, the constraints can be rewritten as

$$\mu_I + \sqrt{\frac{1-\varepsilon}{\varepsilon}}\frac{\sigma_I}{\sqrt{W_I}} \leq \hat{LR} \cdot Pr_I. \tag{8}$$

If we set

$$b_I = \ln\left[\left(\mu_I + \sqrt{\frac{1-\varepsilon}{\varepsilon W_I}}\sigma_I\right)/\hat{LR}\right],$$

we can employ the linear programming formulation (6) for rate estimation. Note that in this case the exposure of each rating cell is incorporated.

5.2. Collective risk model

In the collective risk model, a probability is prescribed for ensuring that the total losses over the whole line of business (LoB) are covered by the premium with a high probability, i.e.

$$P\left(\sum_{I\in\mathcal{I}} L_I^T \leq \sum_{I\in\mathcal{I}} W_I Pr_I\right) \geq 1-\varepsilon.$$

Zaks et al. (2006) proposed the following program for rate estimation, where the mean square error is minimized under the reformulated constraint using the Central Limit Theorem:

$$\begin{aligned} &\min_{Pr_I} \sum_{I\in\mathcal{I}} \frac{1}{r_I}\mathbb{E}\left[(L_I^T - W_I Pr_I)^2\right] \\ &\text{s.t.} \\ &\sum_{I\in\mathcal{I}} W_I Pr_I = \sum_{I\in\mathcal{I}} W_I \mu_I + z_{1-\varepsilon}\sqrt{\sum_{I\in\mathcal{I}} W_I \sigma_I^2}, \end{aligned} \tag{9}$$

where $r_I > 0$ and $z_{1-\varepsilon}$ denotes the quantile of the Normal distribution. Various premium principles can be obtained by the choice of r_I. According to Zaks et al. (2006), Theorem 1, the program has a unique solution

$$\hat{P}r_I = \mu_I + z_{1-\varepsilon}\frac{r_I\sigma}{rW_I},$$

with $r = \sum_{I\in\mathcal{I}} r_I$ and $\sigma^2 = \sum_{I\in\mathcal{I}} W_I\sigma_I^2$. Note that these results were confirmed by Falin (2008) and extended by Frostig et al. (2007). It is possible to use these estimates in the program (6). If we want to incorporate the prescribed loss ratio $\hat{LR}$ for the whole LoB into the previous approach, we can set

$$b_I = \ln\left[\left(\mu_I + z_{1-\varepsilon}\frac{r_I\sigma}{rW_I}\right)/\hat{LR}\right],$$

within the problem (6). Various choices of the weights r_I were discussed by Zaks et al. (2006), e.g. $r_I = 1$ or $r_I = W_I$ were suggested leading to semi-uniform or uniform risk allocations.

6. Numerical example

In this section, we apply the proposed approaches to Motor Third Party Liability contracts. We consider 60 000 policies which are simulated using characteristics of real MTPL portfolio of one of the leading Czech insurance companies. The following indicators are used as the segmentation variables:

- **tariff group**: 5 categories (up to 1000, up to 1350, up to 1850, up to 2500, over 2500 ccm engine),
- **age**: 3 categories (18-30, 30-65, 65 and more years),
- **region**: 4 categories (over 500 000, over 50 000, over 5 000, up to 5 000 inhabitants),
- **gender**: 2 categories (men, women).

Many other available indicators related to a driver (marital status, type of licence), vehicle (engine power, mileage, value), policy (duration, no claim discount) can be used.

We employ the approaches proposed in the previous sections to find the basic premium levels for the tariff groups and the surcharge coefficients for other criteria. The goal loss ratio for new business is set to 0.6 and the maximum feasible surcharge to 100 percent. The parameter estimates for overdipersed Poisson, Gamma and Inverse Gaussian generalized linear models can be found in Table 2. Standard errors and exponentials of the coefficient are also included. All variables are significant based on Wald and likelihood-ratio tests. The parameters of GLM were estimated using SAS GENMOD procedure (SAS/STAT 9.3) and the optimization problems were solved using SAS OPTMODEL procedure (SAS/OR 9.3).

The basic premium levels and surcharge coefficients can be found in Table 3. It is not surprising that the coefficients which are estimated directly from GLM do not fulfill the business requirements and the highest possible surcharge is much higher than 100 percent. This drawback is removed by the optimization problems. The decrease of the surcharge coefficient leads to the increase of the basic premium levels. We refer to the problem where the expected loss is covered as EV model. We used the reliability type model with change constraints and the reformulation based on the Chebyshev's inequality with $\varepsilon = 0.1$, cf. SP model (ind.). Compared with the EV model, the rates increased significantly. This increase is reduced in the second stochastic programming problem based on the collective risk constraint with $\varepsilon = 0.1$, cf. SP model (col.). Note that the stochastic programming models with the Inverse Gaussian distribution of severity lead to higher estimates of the basic premium levels because the estimated variance is much higher then using the Gamma regression.

		Overd. Poisson			Gamma			Inv. Gaussian		
Param.	Level	Est.	Std.Err.	Exp	Est.	Std.Err.	Exp	Est.	Std.Err.	Exp
TG	1	-3.096	0.042	0.045	10.30	0.015	29 778	10.30	0.017	29 765
TG	2	-3.072	0.038	0.046	10.35	0.013	31 357	10.35	0.015	31 380
TG	3	-2.999	0.037	0.050	10.46	0.013	34 913	10.46	0.015	34 928
TG	4	-2.922	0.037	0.054	10.54	0.013	37 801	10.54	0.015	37 814
TG	5	-2.785	0.040	0.062	10.71	0.014	44 666	10.71	0.017	44 679
region	1	0.579	0.033	1.785	0.21	0.014	1.234	0.21	0.016	1.234
region	2	0.460	0.031	1.583	0.11	0.013	1.121	0.11	0.014	1.121
region	3	0.205	0.032	1.228	0.06	0.013	1.059	0.06	0.015	1.058
region	4	0.000	0.000	1.000	0.00	0.000	1.000	0.00	0.000	1.000
age	1	0.431	0.027	1.539	-	-	-	-	-	-
age	2	0.245	0.024	1.277	-	-	-	-	-	-
age	3	0.000	0.000	1.000	-	-	-	-	-	-
gender	1	-0.177	0.018	0.838	-	-	-	-	-	-
gender	2	0.000	0.000	1.000	-	-	-	-	-	-
Scale		0.647	0.000		13.84	0.273		0.002	0.000	

Table 2: Parameter estimates of GLM

		GLM		EV model		SP model (ind.)		SP model (col.)	
		G	IG	G	IG	G	IG	G	IG
TG	1	1 880	1 879	3 805	3 801	9 318	14 952	4 400	5 305
TG	2	2 028	2 029	4 104	4 105	9 979	16 319	8 733	5 563
TG	3	2 430	2 431	4 918	4 918	11 704	19 790	5 547	6 296
TG	4	2 840	2 841	5 748	5 747	13 380	23 145	6 376	7 125
TG	5	3 850	3 851	7 792	7 791	17 453	31 718	8 421	9 169
region	1	2.203	2.201	.311	.390	.407	.552	.463	.407
region	2	.775	.776	.057	.121	.177	.264	.226	.195
region	3	.301	.299	.000	.000	.000	.000	.000	.000
region	4	.000	.000	.000	.000	.000	.000	.000	.000
age	1	.539	.539	.350	.277	.257	.157	.182	.268
age	2	.277	.277	.121	.060	.105	.031	.015	.107
age	3	.000	.000	.000	.000	.000	.000	.000	.000
gender	1	.000	.000	.000	.000	.000	.000	.000	.000
gender	2	.194	.194	.194	.194	.130	.114	.156	.121

Table 3: Estimates of basic premium levels and surcharge coefficients

7. Conclusion

In this paper, we compared several methods for rating of non-life (MTPL) insurance contracts which take into account riskiness of various segments. The probability distribution of losses was described by generalized linear models. Direct application of the estimated coefficient leads to the surcharge coefficients which do not fulfill the business requirements. Therefore, optimization models were introduced. Stochastic programming formulation was employed to consider the distribution of the random losses on a policy.

Future research will be devoted to dynamic models which take into account development of policyholder riskiness. In this case, generalized linear mixed models (Breslow and Clayton 1993) and dynamic stochastic programming models will be employed.

References

[1] Antonio, K., Beirlant, J. (2007) Actuarial statistics with generalized linear mixed models. *Insurance: Mathematics and Economics* 40, 58–76.

[2] Branda, M. (2012a) Chance constrained problems: penalty reformulation and performance of sample approximation technique, *Kybernetika* 48(1), 105–122

[3] Branda, M. (2012b) Sample approximation technique for mixed-integer stochastic programming problems with several chance constraints, *Operations Research Letters* 40(3), 207–211

[4] Branda, M. (2012c) Stochastic programming problems with generalized integrated chance constraints, *Optimization: A Journal of Mathematical Programming and Operations Research* 61(3), 949–968

[5] Branda, M. (2012d) Underwriting risk control in non-life insurance via generalized linear models and stochastic programming. Proceedings of the 30th International Conference on Mathematical Methods in Economics 2012, J. Ramík, D. Stavárek eds., Silesian University in Opava, School of Business Administration in Karviná, 61–66. ISBN: 978-80-7248-779-0

[6] Breslow, N.E., Clayton, D.G. (1993) Approximate inference in generalized linear mixed models. *Journal of the American Statistical Association* 88(421), 9–25.

[7] Bühlmann, H., Gisler, A. (2005) *A course in credibility theory and its applications*. Springer Science & Business.

[8] Cameron, A.C., Trivedi, P.K. (1998) *Regression Analysis of Count Data*, Cambridge: Cambridge University Press

[9] Chen, L., He, S., Zhang, S. (2011) Tight bounds for some risk measures, with applications to robust portfolio selection. *Operations Research* 59(4), 847–865.

[10] Denuit, M., Marchal, X., Pitrebois, S., Walhin, J.-F. (2007) *Actuarial Modelling of Claim Counts: Risk Classification, Credibility and Bonus-Malus Systems*. John Wiley & Sons, Chichester

[11] Ermoliev, Y. M., Ermolieva, T. Y., Macdonald, G. J., Norkin, V. I. (2000) Stochastic optimization of insurance portfolios for managing exposure to catastrophic risks. *Annals of Operations Research* 99, 207–225

[12] Falin, G.I. (2008) On the optimal pricing of a heterogeneous portfolio. *Astin Bulletin* 38(1), 161–170.

[13] Frostig, E., Zaks, Y., Levikson, B. (2007) Optimal pricing for a heterogeneous portfolio for a given risk factor and convex distance measure. *Insurance: Mathematics and Economics* 40(3), 459–67.

[14] Hilli, P., Koivu, M., Pennanen, T. (2011) Cash-flow based valuation of pension liabilities. *European Actuarial Journal* 1(2), 329–343.

[15] de Jong, P., Heller, G.Z. (2008) *Generalized Linear Models for Insurance Data*. Cambridge University Press

[16] McCullagh, P., Nelder, J.A. (1989) *Generalized Linear Models*. 2nd Ed. Chapman and Hall, London

[17] Nemirovski, A., Shapiro, A. (2006) Convex approximations of chance constrained problems. *SIAM Journal on Optimization* 17(4), 969–996

[18] Ohlsson, B. (2008) Combining generalized linear models and credibility models in practice. *Scandinavian Actuarial Journal* 4, 301–314

[19] Ohlsson, J., Johansson, B. (2010) Non-Life Insurance Pricing with Generalized Linear Models. Springer-Verlag Berlin Heidelberg

[20] Prékopa, A. (1995) Stochastic Programming, Kluwer, Dordrecht and Acadmiai Kiad, Budapest

[21] Prékopa, A. (2003) Probabilistic programming, in Stochastic Programming, Handbook in Operations Research and Management Science Vol. 10 (A. Ruszczynski and A. Shapiro, eds.), Elsevier, Amsterdam, 483–554

[22] Withers, Ch., Nadarajah, S. (2011) On the compound Poisson-gamma distribution. *Kybernetika* 47(1), 15–37.

[23] Zaks, Y., Frostig, E., Levikson, B. (2006) Optimal pricing of a heterogeneous portfolio for a given risk level. *Astin Bulletin* 36(1), 161–185.

[24] SAS/STAT 9.3: User's Guide.

[25] SAS/OR 9.3 User's Guide: Mathematical Programming.