Asia-Pacific Journal of Operational Research

OPTIMIZATION APPROACHES TO MULTIPLICATIVE TARIFF OF RATES ESTIMATION IN NON-LIFE INSURANCE

MARTIN BRANDA

Department of Probability and Mathematical Statistics,
Faculty of Mathematics and Physics,
Charles University in Prague,
Sokolovská 83, 186 00 Prague, Czech Republic
branda@karlin.mff.cuni.cz

Received 19 March 2013
Revised 26 January 2014
Accepted 3 April 2014

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 hedging against expected losses and taking into account a prescribed loss ratio and other business requirements. Moreover, we introduce stochastic programming problems with reliability type constraints which take into account individual risk of each rate cell or collective risk. In the numerical study, we apply the approaches to Motor Third Party Liability policies.

Keywords: Non-life insurance; rate making; generalized linear models; optimization models; 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), Frees (2009), Ohlsson (2008), Ohlsson, and Johansson (2010). GLM are used for pure premium estimation based on a priori characteristics of the insurance policy, insured object and policyholder. The frequency–severity approach is the most frequently used, where expected claim count on a policy during one year and expected claim size (severity) 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. We assume that the final rate can be decomposed into a basic premium level and surcharge coefficient which are set according to the a priori characteristics. 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 parts appear. They have already found several applications in insurance, see, e.g., Ermoliev et al. (2000) who managed exposure to catastrophic risks, Consigli et al. (2011) who proposed asset-liability management model for property and casualty insurance, and Hilli et al. (2011) who evaluated pension liabilities. 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 basic approach to pricing in non-life insurance is based on known loss distribution and various principles, e.g. expected value principle, standard deviation, exponential, percentil, see Kaas et al. (2001), Chapter 5 for a review. Zaks et al. (2006) formulated nonlinear programming problems for premium estimation and derived closed formulas for their solutions. These results were confirmed by Falin (2008) who proposed alternative proofs. The optimization problems were extended by Frostig et al. (2007) who used a general distance between the losses and the premium paid. However, these approaches are strongly based on the known distribution of the random losses or at least its first two moments. In this paper, the moments of the distribution are estimated using GLM and the specific rate decomposition is used. Note that pricing of life-insurance contracts is another important task for insurance companies, see, e.g., Bertocchi et al. (2013), Gerber (1997).

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. contract 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 propose basic notation in Section 2. In Section 3, we review definition and basic properties of generalized linear models. We recall a rate-making approach based directly on GLM. In Section 4, we introduce optimization models for rates estimation which enable to take into account various business requirements and the other generalizations proposed above. We extend these models using stochastic programming techniques in Section 5. In Section 6, we apply the proposed methods to Motor Third Party Liability (MTPL) contracts. Section 7 concludes the paper.

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 $i_1 \in \mathcal{I}_1, \dots, i_S \in \mathcal{I}_S$ 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 $I = (i_0, i_1, \dots, i_S)$ with $I \in \mathcal{I} := \mathcal{I}_0 \otimes \mathcal{I}_1 \otimes \cdots \otimes \mathcal{I}_S$. Denote by $\mathcal{I}_{\setminus 0} := \mathcal{I}_1 \otimes \cdots \otimes \mathcal{I}_S$ the index set corresponding to the segmentation criteria only. 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

$$\frac{L_I^T}{TP_I} \leq \hat{LR} \text{ for all } I \in \mathcal{I}. \tag{2.1}$$

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 (2.1) 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

$$\frac{\mathbb{E}[L_I^T]}{TP_I} = \frac{\mathbb{E}[L_I]}{Pr_I} \leq \hat{LR} \text{ for all } I \in \mathcal{I}. \tag{2.2}$$

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 (2.1) are fulfilled with a prescribed probability leading to separate chance (probabilistic) constraints

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

for a small $\varepsilon \in (0, 1)$.

Another approach bounds the loss ratio over the whole line of business:

$$\frac{\sum_{I\in\mathcal{I}} L_I^T}{\sum_{I\in\mathcal{I}} TP_I} \le \hat{LR}. \tag{2.3}$$

Similarly, a probability $1-\varepsilon$ can be prescribed for fulfilling the constraint

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

How the risk is allocated to the 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), cf. Nelder and Wedderburn (1972), 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 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.

Maximum likelihood method is used to estimate the parameters of GLM. Overdispersion is a phenomenon that is often observed in practice for the count data when the variance need not to be equal to the expected value, i.e. standard Poisson

Table 1. Examples of distributions from the exponential family.

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

distribution is not desirable. 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). Zero-inflated models represent an interesting class of regression models for insurance, see Cameron and Trivedi (1998) for an introduction.

3.1. *Pure premium estimation*

The pure premium estimation is based on historical data which contain insurance policies observed from the start date over some period, usually over the first year. We consider a period with reported and settled claims. Indicators related to the contract, insured object and policyholder serve as the predictors in GLM. Note that the time dependent indicators such as car or policy age has to be computed at the start date. New policies are then rated according to the selected rating factors estimated to the policy start date.

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 (2.2). 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 and categorical regressors, then we can get for each $I = (i_0, i_1, \ldots, i_S)$

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

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

$$Pr_{i_0} = \frac{\exp\{\lambda_{i_0} + \gamma_{i_0}\}}{\hat{LR}} \cdot \prod_{s=1}^{S} \min_{i \in \mathcal{I}_s} \exp(\lambda_i) \cdot \prod_{s=1}^{S} \min_{i \in \mathcal{I}_s} \exp(\gamma_i),\ i_0 \in \mathcal{I}_0,$$

$$e_{i_s} = \frac{\exp(\lambda_{i_s})}{\min_{i_s \in I_s} \exp(\lambda_{i_s})} \cdot \frac{\exp(\gamma_{i_s})}{\min_{i_s \in I_s} \exp(\gamma_{i_s})} - 1,\ i_s \in \mathcal{I}_s.$$

Under this choice, the constraints (2.2) 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 employed, interaction or regressors other 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 a natural way 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 random 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 (2.2) with expectation can be rewritten as

$$\mathbb{E}[L_{i_0,i_1,\dots,i_S}] \le \hat{LR} \cdot Pr_{i_0} \cdot (1 + e_{i_1}) \cdot \dots \cdot (1 + e_{i_S}) \text{ for all } I \in \mathcal{I}. \tag{4.4}$$

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. Moreover, the minimization ensures that the surcharge coefficients are set to minimal levels, which are necessary to ensure risk classification. 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}:

$$\begin{aligned} \min \sum_{I\in\mathcal{I}} & w_I Pr_{i_0}(1+e_{i_1})\cdot\dots\cdot(1+e_{i_S}) \\ \hat{LR}\cdot Pr_{i_0}\cdot(1+e_{i_1})\cdot\dots\cdot(1+e_{i_S}) &\geq \mathbb{E}[L_{i_0,i_1,\dots,i_S}],\ \forall(i_0,i_1,\dots,i_S)\in\mathcal{I} \\ (1+e_{i_1})\cdot\dots\cdot(1+e_{i_S}) &\leq 1+r^{max},\ \forall(i_1,\dots,i_S)\in\mathcal{I}_{\setminus 0} \\ e_{i_1},\dots,e_{i_S} &\geq 0,\ \forall(i_1,\dots,i_S)\in\mathcal{I}_{\setminus 0}. \end{aligned} \tag{4.5}$$

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:

$$\begin{aligned} \min \sum_{I\in\mathcal{I}} & w_I e^{u_{i_0}+u_{i_1}+\dots+u_{i_S}} \\ u_{i_0}+u_{i_1}+\dots+u_{i_S} &\geq b_{i_0,i_1,\dots,i_S},\ \forall(i_0,i_1,\dots,i_S)\in\mathcal{I} \\ u_{i_1}+\dots+u_{i_S} &\leq \ln(1+r^{max}),\ \forall(i_1,\dots,i_S)\in\mathcal{I}_{\setminus 0} \\ u_{i_1},\dots,u_{i_S} &\geq 0,\ \forall(i_1,\dots,i_S)\in\mathcal{I}_{\setminus 0}. \end{aligned} \tag{4.6}$$

The problems (4.5) and (4.6) are equivalent in the following sense: $\hat{Pr}_{i_0},\hat{e}_{i_1},\dots,\hat{e}_{i_S}$ is an optimal solution of the problem (4.5) if and only if $\hat{u}_{i_0},\hat{u}_{i_1},\dots,\hat{u}_{i_S}$ is an optimal solution of the problem (4.6) with the relation $\hat{u}_{i_0}=\ln(\hat{Pr}_{i_0})$ and $\hat{u}_{i_s}=\ln(1+\hat{e}_{i_s})$. Note that the estimates do depend on the exposures of the tariff cells in general.

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, we obtain a hardly solvable problem after the logarithmic transform. 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 exactly 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 and not only its expected value as above. We employ chance constraints for satisfying the random individual constraints (2.1) or the collective constraint (2.3) with prescribed levels. However, chance constrained problems are very computationally demanding in general and various approximation methods are usually employed. We refer to Prékopa (1995, 2003), Shapiro et al. (2009) for a comprehensive review of the results and methods available for chance constrained problems. The set of feasible solutions described by chance constraints is usually non-convex. A sufficient condition for convexity is log-concavity of the distribution, cf. Prékopa (1995). Recently, Ninh and Prékopa (2013) proved log-concavity of many compound distributions including Poisson–Gamma distribution. P-level efficient points (pLEPs) introduced by Prékopa (1990) can be employed for solving problems with random right-hand side and discrete distribution. A mathematical programming approach for finding pLEPs was introduced by Lejeune and Noyan (2010). Recently, Lejeune (2012) proposed a new pattern definition of pLEPs and an algorithm based on mixed-integer programming. Sample approximation technique together with a mixed-integer reformulation were investigated by Branda (2012b), Luedtke and Ahmed (2008), who derived an exponential rate of convergence of the approximated solution to the true one. Reliability constraints were discussed by Nemirovski and Shapiro (2006). They proposed a Bernstein approximation for a problem with affine random constraints. Recently, Ji et al. (2013) introduced a new tight convex binary quadratic reformulation for chance-constrained quadratic knapsack problem.

An alternative way to deal with random constraints is to employ penalty functions and to penalize possible violations with respect to the decision vector and random vector simultaneously. The penalized constraints can be incorporated into the objective functions, cf. Branda (2012a, 2013), Ermoliev et al. (2000), or bounded as new constraints leading to (generalized) integrated chance constraints, see Klein Haneveld and van der Vlerk (2006), Branda (2012c). All the mentioned approaches were shown to be asymptotically equivalent under mild conditions by Branda (2012a, 2012c).

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

$$P\left(L^T_{i_0,i_1,\dots,i_S} \leq \hat{LR} \cdot W_{i_0,i_1,\dots,i_S} \cdot Pr_{i_0} \cdot (1+e_{i_1}) \cdot \dots \cdot (1+e_{i_S})\right) \geq 1-\varepsilon,$$

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

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

By setting

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

the formulation (4.6) can be used. However, it can be very difficult and time consuming to compute the quantiles $F^{-1}_{L^T_I}$ for the compound distributions for all $I \in \mathcal{I}$, see, e.g., Withers and Nadarajah (2011), and Central Limit Theorem can not be used, since the expositions W_I of the tariff cells 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

$$P\left(\frac{L^T_I}{TP_I} \geq \hat{LR}\right) \leq \frac{1}{1+(\hat{LR}\cdot TP_I - \mu^T_I)^2/(\sigma^T_I)^2} \leq \varepsilon, \tag{5.7}$$

for $\hat{LR} \cdot TP_I \geq \mu^T_I$. 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.

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

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

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

Note that improved Chebyshev's inequalities were provided by Popescu (2005) for symmetric and symmetric unimodular distributions. However, these estimates are not directly applicable to our problems and some effort into deriving a useful bound will be necessary.

The inequality (5.7) leads to the following constraints, which serve as conservative approximations:

$$\mu^T_I + \sqrt{\frac{1-\varepsilon}{\varepsilon}}\sigma^T_I \leq \hat{LR}\cdot TP_I.$$

Finally, the constraints can be rewritten as

$$\mu_I + \sqrt{\frac{1-\varepsilon}{\varepsilon}} \frac{\sigma_I}{\sqrt{W_I}} \leq \hat{LR} \cdot Pr_I. \tag{5.8}$$

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 (4.6) for rate estimation. Note that in this case the exposure of each rating cell is incorporated directly into b_I. Moreover, when W_I increases, we can easily see that the influence of the standard deviation σ_I on the rate decreases.

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.

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

However, without an additional method we are not able to share the risk surcharges between the policies. Therefore, 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:

$$\begin{aligned} &\min_{Pr_I} \sum_{I \in \mathcal{I}} \frac{1}{r_I} \mathbb{E}\left[(L_I^T - W_I Pr_I)^2 \right] \\ &\text{s.t.} \\ &\sum_{I \in \mathcal{I}} W_I Pr_I = \sum_{I \in \mathcal{I}} W_I \mu_I + z_{1-\varepsilon} \sqrt{\sum_{I \in \mathcal{I}} W_I \sigma_I^2}, \end{aligned} \tag{5.9}$$

where $r_I > 0$ are weights and $z_{1-\varepsilon}$ denotes the quantile of the standard 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

$$\hat{Pr}_I = \mu_I + z_{1-\varepsilon} \frac{r_I \sigma}{r W_I},$$

with $r = \sum_{I \in \mathcal{I}} r_I$ and $\sigma^2 = \sum_{I \in \mathcal{I}} W_I \sigma_I^2$. It is possible to use these estimates in the program (4.6). If we want to incorporate the prescribed loss ratio $\hat{LR}$ for the whole LoB into the previous approach, we can set

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

within the problem (4.6). 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 (MTPL) contracts. We consider policies with settled claims which are simulated using characteristics of a real MTPL portfolio of the leading Czech insurance companies. The basic data contains 60 000 policies, the extended then 600 000. The claim counts were simulated using the Poisson distribution, the corresponding claim severity using the Gamma distribution. The following indicators are used as the regressors and segmentation variables:

- **tariff group**: 5 categories (up to 1000, up to 1350, up to 1850, up to 2500, over 2500 ccm engine),
- **region**: 4 categories (over 500 000, over 50 000, over 5 000, up to 5 000 inhabitants),
- **age**: 3 categories (18-30, 30-65, 65 and more years),
- **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 included into the models and tested for significance.

We employ the approaches proposed in the previous sections to find the basic premium levels for the tariff groups and the surcharge coefficients for the 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. The Inverse Gaussian GLM is employed as an alternative model to the Gamma one. Standard errors and exponentials of the coefficient are also included. All included variables are significant at the 1% level based on the 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).

Since we are using the logarithmic link function, the computation and interpretation of the expected claim numbers and severity can be preformed using the column which contains the exponentials of the estimated parameters. For example, for the expected number of claims for a car with 1400 ccm engine and policyholder-woman, 35 year old, living in a town with 80 000 inhabitants, we obtain

$$\mathbb{E}[N_{3222}] = 0.042 \cdot 1.458 \cdot 1.380 \cdot 1.000 = 0.085.$$

Similarly, for the expected claim severity using the Gamma distribution, we get

$$\mathbb{E}[X_{3222}] = 34252 \cdot 1.116 \cdot 1.000 \cdot 1.000 = 38225.$$

Thus, we can compare directly the contribution of various categories to the riskiness of a contract. The expected number of claims as well as the expected claim severity increase with higher tariff groups based on the engine volume. Similar behaviour can be observed for region where regions with higher number of inhabitants are much riskier. This phenomenon is usually explained by a heavy traffic and expensive cars in such areas. Age and gender of a policyholder have been identified as insignificant in the severity models thus they are omitted. However, younger drivers and men have higher expected number of claims, therefore these factors contribute to the final rates.

The basic premium levels and surcharge coefficients based on GLM, expected value model (EV), individual risk model and collective risk model can be found in Tables 3, 4. The GLM and EV estimates are derived using the basic data, because the simulation technique does not change the results compared with the extended data. 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, in particular 224 percent. This drawback is removed by our optimization approaches. The decrease of the surcharge coefficient leads to the increase of the basic premium levels. We used the reliability type model with change constraints and the reformulation based on the Chebyshev's inequality with $\varepsilon = 0.1$, cf. individual risk model columns. Compared with the EV model, the rates increased significantly. However, higher the exposure is, closer the basic premium levels to the EV ones are. The increase is reduced in the second stochastic programming problem based on the collective risk constraint with the level $\varepsilon = 0.1$, cf. collective risk model columns. Obviously, the price for fulfilling the business requirements is a poorer risk classification for the policies, which is followed by zero coefficients for region 3. Note that the stochastic programming models with the Inverse Gaussian distribution used for severity modeling lead to higher estimates of the basic premium levels because the estimated variance is much higher than using the Gamma regression. This should be used as a warning against using improper model for the severity modeling. No stability in surcharge coefficients is observed, thus the method as well as the exposure influence the estimates.

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, various optimization models were introduced. Stochastic programming formulation was employed to consider the distribution of the random losses on a policy. We showed that the business requirements, selected method, exposure and severity distribution can have a significant influence on the final rates.

Future research will be devoted to dynamic models which can 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.

Table 2. Parameter estimates of GLM (60 000).

Param.	Level	Overd. Poisson Est.	Std.Err.	Exp	Gamma Est.	Std.Err.	Exp	Inv. Gaussian Est.	Std.Err.	Exp
TG	1	-3.446	0.044	0.032	10.31	0.017	30 058	10.31	0.018	30 095
TG	2	-3.293	0.038	0.037	10.36	0.014	31 625	10.36	0.015	31 645
TG	3	-3.181	0.037	0.042	10.44	0.014	34 252	10.44	0.015	34 310
TG	4	-3.126	0.037	0.044	10.53	0.014	37 428	10.53	0.015	37 438
TG	5	-3.019	0.040	0.049	10.70	0.015	44 547	10.71	0.018	44 593
region	1	0.387	0.033	1.473	0.209	0.015	1.233	0.208	0.017	1.231
region	2	0.377	0.030	1.458	0.110	0.013	1.116	0.109	0.015	1.116
region	3	0.103	0.031	1.109	0.066	0.014	1.068	0.064	0.015	1.066
region	4	0.000	0.000	1.000	0.000	0.000	1.000	0.000	0.000	1.000
age	1	0.409	0.028	1.505	-	-	-	-	-	-
age	2	0.322	0.025	1.380	-	-	-	-	-	-
age	3	0.000	0.000	1.000	-	-	-	-	-	-
gender	1	0.172	0.019	1.188	-	-	-	-	-	-
gender	2	0.000	0.000	1.000	-	-	-	-	-	-
Scale	0	0.624	0.000	0.000	12.628	0.264	-	0.002	0.000	-

Table 3. Estimates of basic premium levels and surcharge coefficients for Gamma severity (Exposures in thousands).

Parameter		GLM	EV	Individual risk model Exp. 60	Exp. 300	Exp. 600	Collective risk model Exp. 60	Exp. 300	Exp. 600
TG	1	958	2 590	6 962	4 546	3 973	2 768	2 670	2 646
TG	2	1 175	3 177	8 139	5 396	4 746	3 353	3 256	3 233
TG	3	1 423	3 848	9 531	6 389	5 645	4 023	3 926	3 903
TG	4	1 644	4 445	10 830	7 300	6 464	4 620	4 523	4 500
TG	5	2 176	5 885	13 901	9 470	8 420	6 061	5 964	5 941
region	1	.815	.277	.374	.354	.347	.418	.197	.220
region	2	.628	.146	.236	.217	.211	.282	.077	.097
region	3	.184	.000	.000	.000	.000	.000	.000	.000
region	4	.000	.000	.000	.000	.000	.000	.000	.000
age	1	.505	.318	.295	.292	.289	.203	.415	.386
age	2	.380	.209	.220	.208	.200	.110	.301	.274
age	3	.000	.000	.000	.000	.000	.000	.000	.000
gender	1	.188	.188	.124	.144	.151	.173	.181	.183
gender	2	.000	.000	.000	.000	.000	.000	.000	.000

Acknowledgments

This work was supported by the Czech Science Foundation under the Grant GP13-03749P. I would like to express my gratitude to the anonymous referees, whose comments have greatly improved the paper.

References

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

Table 4. Estimates of basic premium levels and surcharge coefficients for Inverse Gaussian severity (Exposures in thousands).

				Individual risk model			Collective risk model		
		GLM	EV	Exp. 60	Exp. 300	Exp. 600	Exp. 60	Exp. 300	Exp. 600
TG	1	959	2 590	11 883	6 748	5 531	3 030	2 788	2 730
TG	2	1 176	3 176	13 987	8 011	6 594	3 589	3 361	3 307
TG	3	1 425	3 850	16 753	9 621	7 931	4 250	4 029	3 977
TG	4	1 644	4 441	19 555	11 200	9 220	4 841	4 620	4 568
TG	5	2 178	5 885	26 594	15 146	12 434	6 285	6 064	6 011
region	1	.813	.534	.449	.465	.473	.420	.476	.491
region	2	.627	.376	.257	.280	.292	.288	.331	.343
region	3	.182	.000	.000	.000	.000	.000	.000	.000
region	4	.000	.000	.000	.000	.000	.000	.000	.000
age	1	.505	.098	.244	.214	.198	.215	.155	.139
age	2	.380	.007	.180	.143	.124	.125	.063	.047
age	3	.000	.000	.000	.000	.000	.000	.000	.000
gender	1	.188	.188	.109	.125	.133	.159	.174	.178
gender	2	.000	.000	.000	.000	.000	.000	.000	.000

Bertocchi, M, R Giacometti, MC Recchioni and F Zirilli (2013). Pricing life insurance contracts as financial options: the endowment policy case. *Far East Journal of Mathematical Sciences*, Special Volume, Part I, 69–12.

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

Branda, M (2012b). Sample approximation technique for mixed-integer stochastic programming problems with several chance constraints, *Operations Research Letters*, 40(3), 207–21.

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

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

Branda, M (2013). On relations between chance constrained and penalty function problems under discrete distributions. *Mathematical Methods of Operations Research*, 77(2), 265–277.

Breslow, NE and DG Clayton (1993). Approximate inference in generalized linear mixed models. *Journal of the American Statistical Association*, 88(421), 9–25.

Bühlmann, H and A Gisler (2005). *A course in credibility theory and its applications*. Berlin Heidelberg: Springer Science & Business.

Cameron, AC and PK Trivedi (1998). *Regression Analysis of Count Data*. Cambridge: Cambridge University Press.

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

Consigli, G, M Tria, M Gaffo, G Iaquinta, V Moriggia and A Uristani (2011). Dynamic portfolio management for property and casualty insurance. In *Stochastic Optimization Methods in Finance and Energy*, M Bertocchi, G Consigli, MAH Dempster (eds.), 99–124. Springer, New York.

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

Ermoliev, YM, TY Ermolieva, GJ Macdonald and VI Norkin (2000). Stochastic optimization of insurance portfolios for managing exposure to catastrophic risks. *Annals of Operations Research*, 99, 207–225.

Falin, GI (2008). On the optimal pricing of a heterogeneous portfolio. *Astin Bulletin*, 38(1), 161–170.

Frees, EW (2009). *Regression Modeling with Actuarial and Financial Applications.* Cambridge: Cambridge University Press.

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

Gerber, HU (1997). *Life Insurance Mathematics.* Berlin: Springer.

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

Ji, S, X Zheng and X Sun (2013). An improved convex 0-1 quadratic program reformulation for chance-constrained quadratic knapsack problems. *Asia-Pacific Journal of Operational Research*, 30(3), 1340009.

de Jong, P and GZ Heller (2008). *Generalized Linear Models for Insurance Data.* Cambridge: Cambridge University Press.

Kaas, R, M Goovaerts, J Dhaene and M Denuit (2001). *Modern Actuarial Risk Theory.* Dordrech: Kluwer Academic Publishers.

Klein Haneveld, WK and M van der Vlerk (2006). Integrated chance constraints: reduced forms and an algorithm. *Computational Management Science*, 3(4), 245–269.

Lejeune, MA (2012). Pattern definition of the p-efficiency concept. *Annals of Operations Research*, 200, 23–36.

Lejeune, MA and N Noyan (2010). Mathematical programming approaches for generating p-efficient points. *European Journal of Operational Research*, 207(2), 590–600.

Luedtke, J and S Ahmed (2008). A sample approximation approach for optimization with probabilistic constraints. *SIAM Journal on Optimization*, 19, 674–699.

McCullagh, P and JA Nelder (1989). *Generalized Linear Models.* London: Chapman and Hall.

Nelder, J and R Wedderburn (1972). Generalized Linear Models. *Journal of the Royal Statistical Society*, Series A, 135(3), 370–384.

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

Ninh, A and A Prékopa (2013). Log-concavity of compound distributions with applications in stochastic optimization. *Discrete Applied Mathematics*, 161 (18), 3017–3027.

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

Ohlsson, E and B Johansson (2010). *Non-Life Insurance Pricing with Generalized Linear Models.* Berlin Heidelberg: Springer-Verlag.

Popescu, I (2005). A semidefinite programming approach to optimal-moment bounds for convex classes of distributions. *Mathematics of Operations Research*, 30(3), 632–657.

Prékopa, A (1990). Dual method for the solution of a one-stage stochastic programming problem with random RHS obeying a discrete probability distribution. *Mathematical Methods of Operations Research*, 34(6), 441–46.

Prékopa, A (1995). *Stochastic Programming.* Dordrecht: Kluwer, and Budapest: Acadmiai Kiad.

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

Shapiro, A, D Dentcheva and A Ruszczyski (2009). *Lectures on stochastic programming: Modeling and theory*. Philadelphia: SIAM.

Withers, C and S Nadarajah (2011). On the compound Poisson–Gamma distribution. *Kybernetika*, 47(1), 15–37.

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

SAS/STAT 9.3: User's Guide, 2013.

SAS/OR 9.3 User's Guide: Mathematical Programming, 2013.

Martin Branda is Assistant Professor for Financial and Insurance Mathematics at Faculty of Mathematics and Physics of Charles University in Prague since 2011. He got his PhD in Econometric and Operations Research in 2010. His research interests are mathematical programming, stochastic optimization, financial and insurance mathematics. Recently, he has published papers in European Journal of Operational Research, Optimization, Optimization Letters, Operations Research Letters and Annals of Operations Research.