# THE EXPECTED PREMIUM PROPERTIES IN THE BONUS-MALUS SYSTEM

**Risk Management of an Insurance Enterprise: Risk models**

Family and first name: Kryszeń Barbara, PhD
Organization: PZU SA
Al. Jana Pawła II 24
00-133 Warszawa
Poland
Telephone number: +48 22 582 24 69
Fax number: +48 22 582 24 64
E mail: bkryszen@pzu.pl

ABSTRACT

In motor insurance, portfolios are commonly rated both on the base of observed variables and on the base of a bonus-malus system (BMS). The observed variables are called *a priori* variables while a bonus-malus system is the way of an *a posteriori* classification. In the paper the impact of the *a priori* classification on an expected premium in BMS is examined. We introduced two different approaches to the claim frequency modeling in the portfolio: the individual approach and the synthetic approach. In the individual approach the *a priori* classification is taken into consideration, for the synthetic approach – it is not. It means that in the individual approach the hierarchic structure of the mixed distribution is regarded, whereas the synthetic approach takes only the average of the mixed distributions for *a priori* sub-portfolios. The paper shows problems which arise when two classifications are analyzed together, in comparison with the situation where there are treated separately. The effect of the analysis is the classification of approaches based on the level of the portfolios' expected premium.

KEYWORDS

Bonus-malus system, experience rating, claim frequency

## 1. INTRODUCTION

In motor insurance portfolios are commonly rated on the based of observed variables, which values can be determined before the policyholder starts to drive (e.g. sex, type of the car). Policies are partitioned into sub-portfolios with all policyholders belonging to the same sub-portfolio paying the same base premium. The classification variables introduced to partition risks are called *a priori* variables. Despite of the use of many *a priori* variables very heterogeneous behavior can still be observed in each sub-portfolio. Hence in motor insurance the bonus-malus system (BMS) is applied, in which the premium depends on the past claim behavior of a policyholder. This *a posteriori* ratemaking is a way of classifying policyholders according to their risk. For a thorough presentation of the techniques related to BMS see Lamaire (1995).

In the literature BMS is commonly treated as the only means by which premiums are differentiated, so all contracts are subject to the same base premium. Taylor (1997) presented different approach, very closed to the practice. In his paper he addressed the problem of determination of the relative premiums attached to each of the levels of the BMS scale when the *a priori* classification is used by the company.

In our paper the impact of the *a priori* classification on the expected premium in BMS for sub-portfolios and the portfolio is examined. We introduced two different approaches to the claim frequency modeling in the portfolio: the individual approach and the synthetic approach. In the individual approach the *a priori* classification is taken into consideration, for the synthetic approach only the *a posteriori* classification is analyzed. The paper shows problems arise when two classifications are analyzed together, in comparison with the situation where there are treated separately.

The paper is organized as follows. In section 2, we briefly present the modeling for claim frequency in the non-homogenous portfolio, where two approaches are considered: individual and synthetic. Section 3 presents the expected premium properties in sub-portfolios and in the portfolio. Section 4 is devoted to a numerical example, which illustrates the proven theorems. Section 5 provides some final conclusions.

## 2. CLAIM FREQENCY MODELLING IN THE HETEROGENOUS PORTFOILO: INDIVIDUAL AND SYNTHETIC APPROACH

A quintessence of the claim frequency modeling in the heterogeneous portfolio is the allowances for the *a priori* classification and deriving from that different characteristic of the risk in each sub-portfolio. In practice very often one distribution is estimated for the whole portfolio. Taking above into consideration, two different approaches to the claim frequency modeling in the portfolio are introduced: the individual approach and the synthetic approach.

Let $\Theta$ be the random variable describing *a priori* variables in the portfolio and let $H_\Theta(\theta)$ be the distribution for the random variable $\Theta$. The sub-portfolio appointed by the feature $\theta$ is called sub-portfolio $\theta$.

Let $N$ represents the annual number of claims incurred by policyholder. The probability that the policyholder from sub-portfolio $\theta$ has $k$ claims in one year is given by

$$P(N=k \mid \Theta=\theta)=\int_0^\infty p_k(\lambda)dF_\Lambda(\lambda \mid \Theta=\theta) \quad , \tag{2.1}$$

where $p_k(\lambda)=P(N=k \mid \Lambda=\lambda)$ is the probability that a driver with claim frequency $\lambda$ has $k$ claims in one year, the distribution $F_\Lambda(\lambda \mid \Theta=\theta)$ describes claim frequency volatility in sub-portfolio $\theta$. Then the probability that the policyholder from the portfolio has $k$ claims in one year is given by

$$P(N=k)=\int_{\theta\in\Theta}(\int_0^{\infty} p_k(\lambda)dF_{\Lambda}(\lambda\,|\,\Theta=\theta))\,dH_{\Theta}(\theta)\,. \tag{2.2}$$

Formula (2.2) presents the number of claims distribution in the portfolio, where sub-portfolios' characteristics are considered. The presented approach takes into account the differences between *a priori* sub-portfolios. The approach in which the different number of claims distribution for each sub-portfolio is considered is called the individual.

The probability that the policyholder from the portfolio has *k* claims in one year can be also written as:

$$P(N=k)=\int_0^{\infty} p_k(\lambda)\,dG(\lambda)\ . \tag{2.3}$$

The distribution $G(\lambda)$ describes the claims frequency volatility in the whole portfolio. The approach in witch the number of claims distribution is calculated for the whole portfolio is called the synthetic.

The most important difference between those two approaches is that in the individual approach we have to know distributions $F_{\Lambda}(\lambda\,|\,\Theta=\theta)$ and $H_{\Theta}(\theta)$, whereas in the synthetic approach only function $G(\lambda)$ has to be known.

## 3. THE EXPECTED PREMIUM IN THE BONUS-MALUS SYSTEM

Let's start with the short description of BMS. For BMS formally we need the following ingredients:

- $s$ classes numbered by $1,\ldots,s$; the annual premium depends on the number of the actual class;
- a premium scale $b=(b_1,\ldots,b_s)$, where we assumed $b_1\le b_2\le\ldots\le b_s$;
- transition rules which say how to transfer from one class to another is determined once the number of claims is known; once *k* claims are reported, let

  $$t_{ij}^{(k)}=\begin{cases}1 & \text{if the policy gets transferred from class } i \text{ to class } j,\\ 0 & \text{otherwise;}\end{cases}$$
- an initial class $i_0$ for a new policyholder entering the system.

The transition probability, that the policy passed from class *i* to class *j* for the policyholder with claim frequency $\lambda$ is given by

$$p_{ij}(\lambda)=\sum_{k=0}^{\infty} p_k(\lambda)t_{ij}^{(k)}\,. \tag{3.1}$$

Under the assumption that claim frequency of a policyholder is stationary in time, a finite homogeneous Marcov chain with the state space $S=\{1,2,\ldots,s\}$ and transition matrix $M(\lambda)=\lfloor p_{ij}(\lambda)\rfloor$ is a model of BMS (see Lemaire, 1995).

Let $P_n(\lambda)$ be the expected premium in BMS in the year $n\in\mathcal{N}_1$ for the policyholder with the claim frequency $\lambda>0$. The expected premium in the year $n\in\mathcal{N}_1$ can be written as:

$$P_n(\lambda)=d_0\,M^n(\lambda)\,b^T \text{ for } n\in\mathcal{N}_1\,, \tag{3.2}$$

where $d_0=(d_{01},\ldots,d_{0s})$ and $d_{0i}=\begin{cases}1 & i=i_0\\ 0 & i\neq i_0\end{cases}.$

### 3.1. The expected premium for sub-portfolios

In this section we assume that the annual number of claims for the policyholder form the sub-portfolio $\theta \in \Theta$ conforms to the Poisson distribution or any other mixed Poisson distribution. The Poisson distribution is chosen to analyze the homogenous sub-portfolio, the mixed Poisson distribution – to analyze the heterogeneous sub-portfolio.

- Let $P_n^{(\theta)}$ be the expected premium in the year $n \in \mathcal{N}_1$ for the policyholder in the sub-portfolio $\theta$. The premium $P_n^{(\theta)}$ is given by the formula

$$P_n^{(\theta)} = E_F(P_n(\lambda) \mid \theta) = \int_0^\infty P_n(\lambda)\, dF_\Lambda(\lambda | \Theta = \theta). \tag{3.3}$$

- Let $S_n(\theta)$ be the base premium in the year $n \in \mathcal{N}_1$ for the sub-portfolio $\theta$ and let $S_n^{(\theta)}$ be the expected premium in monetary units for the policyholder in the sub-portfolio $\theta$ in the year $n \in \mathcal{N}_1$. The premium $S_n^{(\theta)}$ is given by the formula

$$S_n^{(\theta)} = S_n(\theta) P_n^{(\theta)}. \tag{3.4}$$

In particular, for the homogenous sub-portfolio $\theta$ the annual number of claims in the sub-portfolio conforms to Poisson distribution. Then $F_\Lambda(\lambda \mid \Theta = \theta) = \begin{cases} 0 & \lambda < \overline{\lambda}_\theta \\ 1 & \lambda \geq \overline{\lambda}_\theta \end{cases}$, where $\overline{\lambda}_\theta$ is the average claim frequency in the sub-portfolio $\theta$.

Let ${}_{Poiss}P_n^{(\theta)}$ be the expected premium for the homogenous sub-portfolio $\theta$ in the year $n \in \mathcal{N}_1$. Then

$${}_{Poiss}P_n^{(\theta)} = \int_0^\infty P_n(\lambda) dF_\Lambda(\lambda \mid \Theta = \theta) = P_n(\overline{\lambda}_\theta) \quad \text{for } n \in \mathcal{N}_1. \tag{3.5}$$

Let's consider the situation, where the annual number of claims in a sub-portfolio conforms to a discrete distribution. Let $\lambda_{\theta 1}, \lambda_{\theta 2}, \ldots, \lambda_{\theta k}$, $p_{\theta 1}, p_{\theta 2}, \ldots, p_{\theta k}$ be the parameters of the distribution, where $\lambda_{\theta 1} \leq \lambda_{\theta 2} \leq \ldots \leq \lambda_{\theta k}$ and $\sum_{i=1}^{k} p_{\theta i} = 1$. The discrete distribution is very often used in practice to describe number of claims process in the non-homogenous portfolio (see Walhin, Paris, 1999).

Then let ${}_D P_n^{(\theta)}$ be the expected premium for the sub-portfolio $\theta$ in the year $n \in \mathcal{N}_1$ where the annual number of claims conforms to the discrete distribution. Then

$${}_D P_n^{(\theta)} = \sum_{i=1}^{k} p_{\theta i} P_n(\lambda_{\theta i}) \quad \text{for } n \in \mathcal{N}_1. \tag{3.6}$$

Further in the section we show, that the consequence of taking the sub-portfolio as homogenous, when it should be considered as heterogeneous, is the possibility of underestimating the premium. The purpose of the research is the expression of conditions for which following inequalities are realized:

$$P_n^{(\theta)} \geq {}_D P_n^{(\theta)} \geq {}_{Poiss}P_n^{(\theta)}. \tag{3.7}$$

***Theorem 3.1.*** If the expected premium $P_n(\lambda)$ in BMS for the given year $n \in \mathcal{N}_1$ is the convex function for $\lambda \in (0, +\infty\rangle$, then

$$P_n^{(\theta)} \geq {}_{Poiss}P_n^{(\theta)}.$$

*Proof:* Let's observe that ${}_{Poiss}P_n^{(\theta)} = P_n(E_F(\lambda \mid \theta))$ and $P_n^{(\theta)} = E_F(P_n(\lambda) \mid \theta)$, for any distribution $F_\Lambda(\lambda \mid \Theta = \theta)$ describing claim frequency volatility in sub-portfolio $\theta$.

By the Jensen's inequality for the convex function $P_n(\lambda)$ we have $P_n^{(\theta)} \geq_{Poiss} P_n^{(\theta)}$. □

If the expected premium $P_n(\lambda)$ is the concave function than the opposite inequality is true.

***Theorem 3.2.*** If the expected premium $P_n(\lambda)$ in BMS for the given year $n \in \mathcal{N}_1$ is the convex function for $\lambda \in \langle \varepsilon_1, \varepsilon_2 \rangle$, $0 \leq \varepsilon_1 < \varepsilon_2 \leq +\infty$, and for the parameters $\lambda_{\theta 1}, \lambda_{\theta 2}, \ldots, \lambda_{\theta k}$ of the discrete distribution the following inequality is true $\varepsilon_1 < \lambda_{\theta 1} < \lambda_{\theta 2} \leq \ldots \leq \lambda_{\theta k} \leq \varepsilon_2$, then

$${}_D P_n^{(\theta)} \geq_{Poiss} P_n^{(\theta)}. \quad □$$

The proof of above theorem is based on the Jensen's inequality.

The theorem 3.2 shows that, if parameters $\lambda_{\theta 1}, \lambda_{\theta 2}, \ldots, \lambda_{\theta k}$ of the discrete distribution belong to the interval where $P_n(\lambda)$ is the convex function, then the expected premium calculated under the assumption that sub-portfolio is homogenous, is underestimated. The opposite inequality is true when parameters of the discrete distribution belong to the interval where $P_n(\lambda)$ is the concave function.

In the next theorem two different distributions describing heterogeneity in the sub-portfolio are considered: the continuous distribution $F$ and the discrete distribution with parameters $\lambda_{\theta 1}, \lambda_{\theta 2}, \ldots, \lambda_{\theta k}$, $p_{\theta 1}, p_{\theta 2}, \ldots, p_{\theta k}$. Let's assume that between distributions following relations are valid: for given $0 < \lambda'_{\theta 1} \leq \ldots \leq \lambda'_{\theta(k-1)} < +\infty$ let

$$p_{\theta 1} = \int_0^{\lambda'_{\theta 1}} dF(\lambda), \; p_{\theta i} = \int_{\lambda'_{\theta(i-1)}}^{\lambda'_{\theta i}} dF(\lambda) \text{ for } i = 2,3,\ldots,k-1, \; p_{\theta k} = 1 - \sum_{i=1}^{k-1} p_{\theta i} \text{ and}$$

$$\lambda_{\theta 1} = \frac{1}{p_{\theta 1}} \int_0^{\lambda'_{\theta 1}} \lambda dF(\lambda), \; \lambda_{\theta i} = \frac{1}{p_{\theta i}} \int_{\lambda'_{\theta(i-1)}}^{\lambda'_{\theta i}} \lambda dF(\lambda) \text{ for } i = 2,3,\ldots,k-1, \; \lambda_{\theta k} = \frac{1}{p_{\theta k}} \int_{\lambda'_{\theta(k-1)}}^{+\infty} \lambda dF(\lambda).$$

Then the following theorem is true.

***Theorem 3.3.*** If the expected premium $P_n(\lambda)$ in BMS for the given year $n \in \mathcal{N}_1$ is the convex function for $\lambda \in (0, +\infty\rangle$, then

$$P_n^{(\theta)} \geq_D P_n^{(\theta)}.$$

*Proof:* For the sub-portfolio $\theta$ we have

$$P_n^{(\theta)} = E_F(P_n(\lambda) \mid \theta) = \int_0^{\lambda'_{\theta 1}} P_n(\lambda) dF(\lambda) + \sum_{i=2}^{k-1} \int_{\lambda'_{\theta(i-1)}}^{\lambda'_{\theta i}} P_n(\lambda) dF(\lambda) + \int_{\lambda'_{\theta(k-1)}}^{+\infty} P_n(\lambda) dF(\lambda).$$

Based on the Jensen's inequality we have

$$\int_0^{\lambda'_{\theta 1}} P_n(\lambda) dF(\lambda) = p_{\theta 1} \int_0^{\lambda'_{\theta 1}} P_n(\lambda) \frac{dF(\lambda)}{p_{\theta 1}} \geq p_{\theta 1} P_n(\lambda_{\theta 1})$$

and

$$\int_{\lambda'_{\theta(i-1)}}^{\lambda'_{\theta i}} P_n(\lambda) dF(\lambda) = p_{\theta i} \int_{\lambda'_{\theta(i-1)}}^{\lambda'_{\theta i}} P_n(\lambda) \frac{dF(\lambda)}{p_{\theta i}} \geq p_{\theta i} P_n(\lambda_{\theta i}) \text{ for } i = 2,3,\ldots,k-1,$$

$$\int_{\lambda'_{\theta(k-1)}}^{+\infty} P_n(\lambda) dF(\lambda) = p_{\theta k} \int_{\lambda'_{\theta(k-1)}}^{+\infty} P_n(\lambda) \frac{dF(\lambda)}{p_{\theta k}} \geq p_{\theta k} P_n(\lambda_{\theta k}).$$

By the Jensen's inequality for the convex function $P_n(\lambda)$ the inequality $E_F(P_n(\lambda) \mid \theta) \geq \sum_{i=1}^{k} p_{\theta i} P_n(\lambda_{\theta i})$ is valid. Then $P_n^{(\theta)} \geq_D P_n^{(\theta)}$ as claimed. □

Theorems 3.1-3.3 describe the situation where the function $P_n(\lambda)$ is convex (concave) for the whole interval of $\lambda$. In practice the shape of the premium depends on the BMS constructions, usually it changes every year. The following theorem considers situation which is very common for the BMS described in the numerical example (see section 4).

***Lemma 3.1.*** Let $\varphi(\lambda)$ for $\lambda > 0$ be a non-linear function, continuous, with positive values, convex for $\lambda \in (0,\varepsilon\rangle$ and concave for $\lambda \in \langle\varepsilon, L)$. Additionally, let $\lambda_1$ and $\lambda_2$ be the given numbers so, that $\lambda_1 \in (0,\varepsilon\rangle$, $\lambda_2 \in \langle\varepsilon, L)$ and $\lambda_1 \neq \lambda_2$. Then only one from the following three relations between $p\varphi(\lambda_1)+(1-p)\varphi(\lambda_2)$ and $\varphi(p\lambda_1+(1-p)\lambda_2)$, where $p \in (0,1)$, is true:

(A) $\underset{p^* \in (0,1)}{\exists} p^*\varphi(\lambda_1)+(1-p^*)\varphi(\lambda_2) = \varphi(p^*\lambda_1+(1-p^*)\lambda_2)$

and then

$\underset{p \in (p^*,1)}{\forall} p\varphi(\lambda_1)+(1-p)\varphi(\lambda_2) > \varphi(p\lambda_1+(1-p)\lambda_2)$,

$\underset{p \in (0,p^*)}{\forall} p\varphi(\lambda_1)+(1-p)\varphi(\lambda_2) < \varphi(p\lambda_1+(1-p)\lambda_2)$,

(B) $\underset{p \in (0,1)}{\forall} p\varphi(\lambda_1)+(1-p)\varphi(\lambda_2) > \varphi(p\lambda_1+(1-p)\lambda_2)$,

(C) $\underset{p \in (0,1)}{\forall} p\varphi(\lambda_1)+(1-p)\varphi(\lambda_2) < \varphi(p\lambda_1+(1-p)\lambda_2)$. □

Because of the complication of the calculations, the theorem 3.4 is devoted only to the situation, where the annual number of claims conforms to the discrete distribution with parameters $\lambda_{\theta 1}, \lambda_{\theta 2}, p_{\theta 1}, p_{\theta 2} = 1 - p_{\theta 1}$.

***Theorem 3.4.*** Let $P_n(\lambda)$ be the convex function for $\lambda \in (0,\varepsilon\rangle$ and concave function for $\lambda \in \langle\varepsilon, L)$, where $0 < \varepsilon < L$, and for parameters $\lambda_{\theta 1}, \lambda_{\theta 2}$ of the discrete distribution in the sub-portfolio $\theta$ following relation is true $0 < \lambda_{\theta 1} \le \varepsilon \le \lambda_{\theta 2} < L$. The necessary and sufficient condition for the inequality to be true

$${}_D P_n^{(\theta)} \ge {}_{Poiss} P_n^{(\theta)} \tag{3.8}$$

is the existence of the solution $p = p_n^* \in (0, p_{\theta 1}\rangle$ of the equation

$$pP_n(\lambda_{\theta 1}) + (1-p)P_n(\lambda_{\theta 2}) = P_n(p\lambda_{\theta 1} + (1-p)\lambda_{\theta 2}) \tag{3.9}$$

or, when the equation (3.9) has no solution, that for every number $p \in (0,1)$ the following inequality is valid

$$pP_n(\lambda_{\theta 1}) + (1-p)P_n(\lambda_{\theta 2}) > P_n(p\lambda_{\theta 1} + (1-p)\lambda_{\theta 2}).$$

*Proof:* Based on the lemma 3.1 for the function $P_n(\lambda)$ only one from the following three conditions is true:

(A) $\underset{p \in (0,1)}{\exists} pP_n(\lambda_{\theta 1}) + (1-p)P_n(\lambda_{\theta 2}) = P_n(p\lambda_{\theta 1} + (1-p)\lambda_{\theta 2})$,

(B) $\underset{p \in (0,1)}{\forall} pP_n(\lambda_{\theta 1}) + (1-p)P_n(\lambda_{\theta 2}) > P_n(p\lambda_{\theta 1} + (1-p)\lambda_{\theta 2})$,

(C) $\underset{p \in (0,1)}{\forall} pP_n(\lambda_{\theta 1}) + (1-p)P_n(\lambda_{\theta 2}) < P_n(p\lambda_{\theta 1} + (1-p)\lambda_{\theta 2})$.

If (A) is true, so when the equation (3.9) has the solution $p \in (0,1)$, then the following two situations are possible:

- $p \in (0, p_{\theta 1}\rangle$ and then based on the lemma 3.1 the following inequality is valid
  $p_{\theta 1}P_n(\lambda_{\theta 1}) + (1-p_{\theta 1})P_n(\lambda_{\theta 2}) > P_n(p_{\theta 1}\lambda_{\theta 1} + (1-p_{\theta 1})\lambda_{\theta 2})$,
  and then ten inequality (3.8) is true.
- $p \in (p_{\theta 1}, 1)$ and then based on the lemma 3.1 the following inequality is valid
  $p_{\theta 1}P_n(\lambda_{\theta 1}) + (1-p_{\theta 1})P_n(\lambda_{\theta 2}) < P_n(p_{\theta 1}\lambda_{\theta 1} + (1-p_{\theta 1})\lambda_{\theta 2})$,
  and then ten inequality (3.8) is not true.

In the situation that equation (3.9) has no solution in the interval (0,1), then only one from the conditions (B) or (C) is valid. For the condition (B) the following inequality is true

$$p_{\theta 1}P_n(\lambda_{\theta 1}) + (1-p_{\theta 1})P_n(\lambda_{\theta 2}) > P_n(p_{\theta 1}\lambda_{\theta 1} + (1-p_{\theta 1})\lambda_{\theta 2}),$$

and then the inequality (3.8) is valid. For the condition (C) the following inequality is true
$p_{\theta 1}P_n(\lambda_{\theta 1})+(1-p_{\theta 1})P_n(\lambda_{\theta 2})<P_n(p_{\theta 1}\lambda_{\theta 1}+(1-p_{\theta 1})\lambda_{\theta 2})$,
and then the inequality (3.8) is not valid. □

### 3.2.The expected premium for the portfolio

For the analyze of the expected premium's properties in the portfolio we introduced the following notation.

- Let ${}_I S_n$ be the expected premium in the monetary units in the portfolio for the year $n\in\mathcal{N}_1$ for the individual approach. The premium ${}_I S_n$ can be written as

$$ {}_I S_n = \int_{\theta\in\Theta} S_n(\theta)\Big(\int_0^\infty P_n(\lambda)\,dF_\Lambda(\lambda\mid\Theta=\theta)\Big)dH_\Theta(\theta) = \int_{\theta\in\Theta} S_n(\theta)P_n^{(\theta)}\,dH_\Theta(\theta)\ . \tag{3.10}$$

For the synthetic approach the *a priori* classification is not taken into account. Following that we assume that all policyholders are subject to the same base premium.

- Let $S_n$ be the base premium in the year $n\in\mathcal{N}_1$ in the portfolio for the synthetic approach, and let ${}_S S_n$ be the expected premium in the monetary units in the portfolio for the year $n\in\mathcal{N}_1$ for the synthetic approach. The premium ${}_S S_n$ can be written as

$$ {}_S S_n = S_n\int_0^\infty P_n(\lambda)\,dG(\lambda)\ . \tag{3.11}$$

Let's consider the portfolio divided into *l* sub-portfolios, where $l>2$, according to the *a priori* classification. That portfolio is later called D-portfolio. Let $\overline{\lambda}$ be the average claim frequency in the portfolio. For each $\theta\in\{1,2,\ldots l\}$ let $\overline{\lambda}_\theta$ be the average claim frequency in sub-portfolio *θ*. Additionally, let $p_\theta$ be the share of the sub-portfolio *θ* in the whole portfolio, then

$$h_\Theta(\theta)=p_\theta,\ \text{where}\ p_\theta>0,\ \sum_{\theta=1}^{l}p_\theta=1,\ \text{for}\ \theta\in\{1,2,\ldots,l\}\ \text{and}\ \overline{\lambda}=\sum_{\theta=1}^{l}p_\theta\overline{\lambda}_\theta\ .$$

For each sub-portfolio claim frequency distributions are $F_\Lambda(\lambda|\Theta=\theta)$ for $\theta\in\{1,2,\ldots,l\}$.
The following theorems have very important meaning for practice, because every real portfolio divided into sub-portfolios according to *a priori* variables can be considered as D-portfolio.
***Theorem 3.5.*** If the portfolio is the D-portfolio, then:

1. for the synthetic approach the expected premium in the portfolio for the year $n\in\mathcal{N}_1$ can be written as:

$$ {}_S S_n = S_n\sum_{\theta=1}^{l}p_\theta P_n(\overline{\lambda}_\theta)\,; \tag{3.12}$$

2. for the individual approach to the homogenous sub-portfolios the expected premium in the portfolio for the year $n\in\mathcal{N}_1$ can be written as:

$$ {}_{I,Poiss} S_n = \sum_{\theta=1}^{l}p_\theta S_n(\theta)\,{}_{Poiss}P_n^{(\theta)}\ ; \tag{3.13}$$

3. for the individual approach to the heterogeneous sub-portfolios the expected premium in the portfolio for the year $n\in\mathcal{N}_1$ can be written as:

$$ {}_I S_n = \sum_{\theta=1}^{l}p_\theta S_n(\theta)P_n^{(\theta)}\ . \tag{3.14}$$

□

An object of the analysis is the thesis, that the consequence of the usage of the synthetic approaches instead the individual approach to homogeneous or heterogeneous sub-portfolios, is the possibility of underestimating the expected premium. In the formal sense the main purpose is to express the conditions under which following inequalities are realized:

$${}_{I}S_n \geq {}_{I,Poiss}S_n \geq {}_{S}S_n .$$

***Theorem 3.6.*** In the given year $n \in \mathcal{N}_1$ for the expected premium in the D-portfolio the sufficient condition for the following inequality to be valid

$${}_{I}S_n \geq {}_{I,Poiss}S_n$$

is, that for each sub-portfolio $\theta \in \{1,2\ldots,l\}$ in the year *n* the following inequality is true

$$P_n^{(\theta)} \geq_{Poiss} P_n^{(\theta)} .$$

□

***Theorem 3.7.*** If for D-portfolio in the given year $n \in \mathcal{N}_1$ the expected premium $P_n(\lambda)$ is a non-decreasing function of $\lambda$, and base premiums follow the relations $S_n = \gamma_n \overline{\lambda}$ and $S_n(\theta) = \gamma_n \overline{\lambda}_\theta$ for $\theta \in \{1,2\ldots,l\}$, $\gamma(n) \geq 1$, then the following inequality is valid

$${}_{I,Poiss}S_n \geq {}_{S}S_n .$$

*Proof:* In D-portfolio the following relation is true

$$\overline{\lambda} = \sum_{\theta=1}^{l} p_\theta \overline{\lambda}_\theta .$$

From the formula (3.13) we have

$${}_{I,Poiss}S_n = \gamma_n \sum_{\theta=1}^{l} p_\theta \overline{\lambda}_\theta P_n(\overline{\lambda}_\theta) ,$$

And from the formula (3.12), we have

$${}_{S}S_n = \gamma_n \overline{\lambda} \sum_{\theta=1}^{l} p_\theta P_n(\overline{\lambda}_\theta) .$$

Let’s observe that

$$\sum_{\theta=1}^{l} p_\theta \left(\overline{\lambda} - \overline{\lambda}_\theta\right) = \sum_{\theta=1}^{l} p_\theta \overline{\lambda} - \sum_{\theta=1}^{l} p_\theta \overline{\lambda}_\theta = \overline{\lambda} - \overline{\lambda} = 0 .$$

The expected premium $P_n(\lambda)$ is a non-decreasing function of $\lambda > 0$, so following conditions are true:

- $\underset{n \in \mathcal{N}_1}{\forall} [\overline{\lambda}_\theta - \overline{\lambda} \leq 0 \Rightarrow P_n(\overline{\lambda}_\theta) \leq P_n(\overline{\lambda})]$ and
- $\underset{n \in \mathcal{N}_1}{\forall} [\overline{\lambda}_\theta - \overline{\lambda} \geq 0 \Rightarrow P_n(\overline{\lambda}_\theta) \geq P_n(\overline{\lambda})]$.

Based on the above we have

$${}_{I,Poiss}S_n - {}_{S}S_n = \gamma_n \sum_{\theta=1}^{l} p_\theta P_n(\overline{\lambda}_\theta)(\overline{\lambda}_\theta - \overline{\lambda}) =$$

$$= \gamma_n \left[ \sum_{\overline{\lambda}_\theta < \overline{\lambda}} P_n(\overline{\lambda}_\theta) p_\theta (\overline{\lambda}_\theta - \overline{\lambda}) + \sum_{\overline{\lambda}_\theta \geq \overline{\lambda}} P_n(\overline{\lambda}_\theta) p_\theta (\overline{\lambda}_\theta - \overline{\lambda}) \right] =$$

$$\geq \gamma_n \left[ \sum_{\overline{\lambda}_\theta < \overline{\lambda}} P_n(\overline{\lambda}) p_\theta (\overline{\lambda}_\theta - \overline{\lambda}) + \sum_{\overline{\lambda}_\theta \geq \overline{\lambda}} P_n(\overline{\lambda}) p_\theta (\overline{\lambda}_\theta - \overline{\lambda}) \right] = \gamma_n P_n(\overline{\lambda}) \sum_{\theta=1}^{l} p_\theta (\overline{\lambda}_\theta - \overline{\lambda}) = 0,$$

so

${}_{I,Poiss}S_n \geq {}_{S}S_n$, as claimed. □

***Corollary 3.1.*** If assumptions of the theorem 3.6 and the theorem 3.7 are valid, then following inequalities are true

$${}_{I}S_n \geq {}_{I,Poiss}S_n \geq {}_{S}S_n . \quad (3.15)$$

## 4. NUMERICAL EXAMPLE

In order to illustrate the application of the proven theorems, we analyze the BMS currently applied in motor third-party liability insurance by PZU SA, the Polish insurance company. The PZU BMS consists of 13 classes. New policyholder enters the system in class 5. The specific premium level for each class as well as transition rules are provided in Table 1.

Table 1. BMS of PZU SA

| Class number | Premium level (in percentage) | Class number after | | | | | | |
|---|---|---|---|---|---|---|---|---|
| | | 0 | 1 | 2 | 3 | 4 | 5 | 6 or more |
| | | claims | | | | | | |
| 1 | 200 | 2 | 1 | 1 | 1 | 1 | 1 | 1 |
| 2 | 150 | 3 | 1 | 1 | 1 | 1 | 1 | 1 |
| 3 | 130 | 4 | 1 | 1 | 1 | 1 | 1 | 1 |
| 4 | 115 | 5 | 2 | 1 | 1 | 1 | 1 | 1 |
| 5 | 100 | 6 | 3 | 1 | 1 | 1 | 1 | 1 |
| 6 | 90 | 7 | 4 | 2 | 1 | 1 | 1 | 1 |
| 7 | 80 | 8 | 5 | 3 | 1 | 1 | 1 | 1 |
| 8 | 80 | 9 | 6 | 4 | 2 | 1 | 1 | 1 |
| 9 | 70 | 10 | 7 | 5 | 3 | 1 | 1 | 1 |
| 10 | 60 | 11 | 8 | 6 | 4 | 2 | 1 | 1 |
| 11 | 50 | 12 | 9 | 7 | 5 | 3 | 1 | 1 |
| 12 | 50 | 13 | 10 | 8 | 6 | 4 | 2 | 1 |
| 13 | 40 | 13 | 11 | 9 | 7 | 5 | 3 | 1 |

For the PZU BMS the numerical calculations were conducted in order to present intervals where the expected premium $P_n(\lambda)$ is the convex function and intervals where it is the concave function for $\lambda \in (0,1)$.

Table 2. Convex intervals and concave intervals for the expected premium in PZU BMS: $P_n(\lambda)$ for $n$=1,2,…,10

| The expected premium | Convex intervals | Concave intervals |
|---|---|---|
| $P_1(\lambda)$ | (0,0.43) | (0.43,1) |
| $P_2(\lambda)$ | (0,0.27) | (0.27,1) |
| $P_3(\lambda)$ | (0,0.36) | (0.36;1) |

| | | |
|---|---|---|
| $P_4(\lambda)$ | (0,0.30) | (0.30;1) |
| $P_5(\lambda)$ | (0,0.27) | (0.27,1) |
| $P_6(\lambda)$ | (0.13,0.21) | (0,0.13) and (0.21,1) |
| $P_7(\lambda)$ | (0.01,0.28) | (0,0.01) and (0.28,1) |
| $P_8(\lambda)$ | (0.01,0.26) | (0,0.01) and (0.26,1) |
| $P_9(\lambda)$ | (0,0.25) | (0.25;1) |
| $P_{10}(\lambda)$ | (0,0.28) | (0.28;1) |

For the chosen PZU's sub-portfolios we calculated the expected premium for years $n = 1,2,\ldots,10$. The values of parameters for the number of claims distribution for each sub-portfolio were taken on the bases of the PZU's empirical data. Table 3 presents the values of the expected premium ${}_{Poiss}P_n^{(\theta)}$, ${}_{D}P_n^{(\theta)}$ in sub-portfolios A and B. In case where the inequality ${}_{D}P_n^{(\theta)} > {}_{Poiss}P_n^{(\theta)}$ is not true, the values are marked by cursive.

Table 3. The expected premium in sub-portfolios A and B

| Period | Sub-portfolio A | | Sub-portfolio B | |
|---|---|---|---|---|
| | ${}_{Poiss}P_n$ | ${}_{D}P_n$ | ${}_{Poiss}P_n$ | ${}_{D}P_n$ |
| *n*=1 | 92,4454 | 92,5773 | 91,9469 | 92,0500 |
| *n*=2 | 84,3525 | 84,4497 | 83,4553 | 83,5964 |
| *n*=3 | 83,9442 | 84,6166 | 83,0993 | 83,6839 |
| *n*=4 | 75,2425 | 75,8472 | 74,1200 | 74,7502 |
| *n*=5 | 66,5240 | 66,9129 | 65,1299 | 65,6959 |
| *n*=6 | *60,2306* | *58,9217* | *58,2420* | *57,5335* |
| *n*=7 | *58,4212* | *58,3246* | 56,7068 | 56,9800 |
| *n*=8 | *49,7154* | *49,2730* | 47,7285 | 47,8369 |
| *n*=9 | 47,5047 | 48,2267 | 45,7245 | 46,8580 |
| *n*=10 | 47,3475 | 48,3533 | 45,6809 | 46,9910 |

When we compare the results shown in table 3 with intervals in which the expected premium of the PZU BMS is convex (table 2), then it can be easily seen, that in all cases where the function $P_n(\lambda)$ is convex in the right-side neighborhood of zero, ${}_{D}P_n^{(\theta)}$ is greater than ${}_{Poiss}P_n^{(\theta)}$.

As the example of the D-portfolio we considered the sub-portfolio, which groups policyholders from sub-portfolio A and sub-portfolio B. As the result we get portfolio, later called portfolio AB. The values of parameters for the number of claims distribution for the portfolio AB were taken based on the distributions estimated for each sub-portfolio A and B. We assume that for base premiums following relations are true: $S_n = \gamma_n \bar{\lambda}$ and $S_n(\theta) = \gamma_n \bar{\lambda}_\theta$, and for the simplicity we take $\gamma_n = 1$. Table 4 presents the amounts of the expected premium for the portfolio AB ${}_{S}S_n$, ${}_{I,Poiss}S_n$, ${}_{I}S_n$ for $n = 1,2,\ldots,10$. In case where the inequality (3.15) is not true, the values are marked by cursive.

Table 4. The expected premium in portfolio AB

| Period | Portfolio AB | | |
|---|---|---|---|
| | ${}_{S}S_{n}$ | ${}_{I,Poiss}S_{n}$ | ${}_{I}S_{n}$ |
| *n*=1 | 4,6767 | 4,6778 | 4,6835 |
| *n*=2 | 4,2503 | 4,2523 | 4,2588 |
| *n*=3 | 4,2316 | 4,2334 | 4,2644 |
| *n*=4 | 3,7789 | 3,7814 | 3,8130 |
| *n*=5 | 3,3256 | 3,3287 | 3,3549 |
| *n*=6 | 2,9831 | *2,9875* | *2,9428* |
| *n*=7 | 2,9017 | 2,9055 | 2,9140 |
| *n*=8 | 2,4491 | *2,4535* | *2,4509* |
| *n*=9 | 2,3447 | 2,3487 | 2,4002 |
| *n*=10 | 2,3411 | 2,3448 | 2,4069 |

When we compare the above results with intervals in which the expected premium of the PZU BMS is convex (table 2), then it can be easily seen, that in all cases where the function $P_n(\lambda)$ is convex in the right-side neighborhood of zero, the inequality (3.15) is valid. The inequality (3.15) is not true for the years 6 and 8, because for these years assumptions of the theorem 3.6 are not valid.

For the insurance company a consequence of the usage of the synthetic approach or the individual approach to the homogenous portfolio is the possibility of underestimating the premium in comparison with the individual approach to heterogeneous sub-portfolios. The greatest mistake we receive for synthetic model, wherein the *a priori* classification is not taken into account. Besides, the individual approach to heterogeneous sub-portfolio is the closest to the portfolio characteristics, one ought to find it most adequate.

## 5. CONCLUSIONS

Received dependences indicate the possibility of the commission of errors in the premium valuation for the whole portfolio, if unsuitable assumptions are taken. A result of the research show how important is to regard risk heterogeneity in sub-portfolios. Additionally, in the portfolio rated on the based of *a priori* variables, policyholders classified to different sub-portfolios, pay the base premium in different height, depending on the sub-portfolio characteristics. Only the individual approach to heterogeneous sub-portfolios satisfies postulates of the correct premium calculation. In this approach different level of the *a priori* premium and different risk characteristic in each sub-portfolio are considered. For a premium calculation it is necessary to regards both, the *a priori* classification and the different characteristics of the risk in sub-portfolios.

In practice these results can be applied to asses the potential risk connected with simplified approach. This is particularly important for insurance companies due to the fact that it allows adjusting the bonus-malus system construction to the portfolio characteristics as well as to project future cash flow.

REFERENCES

LEMAIRE, J. (1995) *Bonus-Malus System in Automobile Insurance.* Kluwer Academic Publisher, Boston.

PITREBOIS. S., DENUIT, M. and WALHIN, J.F. (2003) *Setting a Bonus-Malus scale in the presence of other rating factors: Taylor's work revisited, ASTIN Bulletin* **33**, 419-336.

ROLSKI, T., SCHMIDLI, H., SCHMIDT, V. and TEUGELS, J. (1999) *Stochastic Processes for Insurance and Finance.* John Wiley & Sons, New York.

TAYLOR, G. (1997) *Setting a Bonus-Malus scale in the presence of other rating factors, ASTIN Bulletin* **27**, 319-327.

WALHIN J.F., PARIS J. (1999) *Using Mixed Poisson Processes in Connection with Bonus-malus System*, *ASTIN Bulletin* **29**, 81-99.