ACTA UNIVERSITATIS LODZIENSIS
FOLIA OECONOMICA 196, 2006

*Agata Boratyńska**

ROBUST BAYESIAN PREDICTION WITH ASYMMETRIC LOSS FUNCTION IN POISSON MODEL OF INSURANCE RISK

Abstract. In robust Bayesian analysis a prior is assumed to belong to a family instead of being specified exactly. The multiplicity of priors leads to a collection of Bayes actions. It is clearly essential to be able to recommend one action (estimate, predictor) from this set.

We consider the problem of robust Bayesian prediction of a Poisson random variable under LINEX loss. Some uncertainty about the prior is assumed by introducing three classes of conjugate priors. The conditional Γ-minimax predictors and posterior regret Γ-minimax predictors are constructed. The application to the collective risk model is presented.

Key words: Bayesian prediction, Bayesian robustness, LINEX loss, family of priors, collective risk model.

1. INTRODUCTION

We consider a Bayesian risk model. Our objective is to predict the number of future claims in order to calculate an appropriate premium. We will use the Bayesian forecasting, which combines the knowledge about characteristics in the whole portfolio with knowledge about an individual contract. The knowledge about an individual is given in the form of a random sample $X = (X_1, X_2, ..., X_n)$. The probability distribution of its sample and a predicted random variable Y depends on an unknown parameter (characteristic) θ. The knowledge about the whole portfolio is presented by using a prior distribution of θ.

The standard Bayesian analysis in the risk theory has been considered in many papers, for examples see Makov et al. (1996) and Klugman et al. (1998). In classical credibility (cf. Gooverts 1990, Klugman 1992) we find

* Ph.D., Institute of Econometrics, Warsaw School of Economics.

the predictor $\hat{g}(X)$ of random variable Y by minimizing the mean square error i.e.

$$E(Y-\hat{g}(X))^2,$$

where the operator E emphasizes the expectation with respect to the joint probability distribution of all random variables θ, X, Y. For square loss function positive and negative deviations have the same weight. In insurance the penalty for underestimation should not be the same as for overestimation. In contrast with the classical results we choose an asymmetric loss function, LINEX loss function (cf. Zellner 1986), equal

$$L(y,d)=e^{c(y-d)}-c(y-d)-1,$$

where $c \neq 0$ is a known parameter and y is a value of a predicted random variable and d is a value of a predictor. If $c>0$ then underestimation gives greater error than overestimation (for $c<0$ overestimation gives greater error than underestimation). Without loss of generality we will assume $c>0$. The LINEX loss is connected with the premium calculated according to the well-known "exponential principle", (cf. Goovaerts 1990). Note that the square loss gives the net premium.

The obtained predictor depends on a choice of a prior distribution Π. In most Bayesian analysis the elicitation of a prior is quite difficult and can be uncertain. The robust Bayesian inference uses a class Γ of prior distributions, which model uncertainty of the prior information. It deals with the problem of measuring the range of a posterior quantity (for example: the range of a Bayes estimator, a predictor, a posterior risk) while a prior distribution runs over the class Γ. Its aim is also finding robust procedures. The general references on robust Bayesian methods are Berger (1990, 1994), Insua and Ruggeri (2000). In insurance models the range of a premium, when priors run over a class Γ, has been considered in many papers, for examples see Insua et al. (1999), Gómez-Déniz et al. (1999, 2002), and references therein.

We introduce two measures of robustness of a predictor and find the optimal solutions (the most robust predictors of number of future claims) with respect to these measures. The next section presents a general definition of a Bayes predictor under LINEX loss and definitions of robust predictors: the conditional Γ-minimax predictor and the posterior regret Γ-minimax predictor. Section 3 gives the robust predictors in the Poisson model with three classes of conjugate priors and presents the application in the collective risk model.

2. BAYESIAN PREDICTION UNDER LINEX LOSS

Let X be an observed random variable with a probability distribution P_θ indexed by a real parameter θ, with a density p_θ with respect to some σ-finite measure. Let Y be a real random variable with a probability distribution F_θ indexed the same parameter θ with density f_θ with respect to the σ-finite measure μ. Variables X and Y are conditionally independent given θ. Suppose that θ has a prior distribution Π.

Our goal is to predict Y under LINEX loss. We are to find a Bayes predictor, i.e. a statistic $\hat{g}^{\Pi}(X)$, which minimizes the quantity $E_\Pi L(Y, \hat{g}(X))$ with respect to $\hat{g}$. The subscript Π in the operator E_Π denotes that the prior Π is and $E_\Pi L(Y, \hat{g}(X))$ is an expected value of a function L with respect to the joint probability distribution of θ, X and Y.
We have

$$E_\Pi L(Y, \hat{g}(X)) = E_\Pi \mathrm{E}_\Pi [L(Y, \hat{g}(X)) \,|\, X].$$

To find the Bayes predictor for $X = x$ it is enough to minimize the posterior risk

$$R_x(\Pi, d) = E_\Pi[L(Y, d)\,|\,x] = e^{-cd} \mathrm{E}_\Pi(e^{cY}|x) + cd - cE_\Pi(Y|x) - 1$$

over $d \in \mathfrak{R}$. Thus the Bayes predictor at a point x is

$$\hat{g}^{\Pi}(x) = \frac{1}{c} \ln E_\Pi(\mathrm{e}^{cY}|x).$$

Let $\Pi(\cdot|x)$ denote the posterior distribution of θ if $X = x$. Then

$$\hat{g}^{\Pi}(x) = \frac{1}{c} \ln \int_{\Theta} \int_{R} e^{cy} f_\theta(y) \mu dy) \Pi(d\theta|x) = -\frac{1}{c} \ln E_\Pi(M_Y(c\,|\,\theta)\,|\,x),$$

where $M_Y(t|\theta)$ denotes the moment generation function of the random variable Y at a point t given θ, and $E_\Pi(h(\theta)|x)$ denotes the expected value of a function h, when θ has a posterior distribution.

Now assume that a prior Π is not specified exactly and consider a class Γ of priors. Consider two functions as two measures of robustness of a predictor $\hat{g}$ at the point x:

- $(x, \hat{g}(x)) \longrightarrow \sup_{\Pi \in \Gamma} R_x(\Pi . \hat{g}(x))$;
- $(x, \hat{g}(x)) \longrightarrow \sup_{\Pi \in \Gamma} U_x(\Pi, \hat{g}(x))$, where

$$U_x(\Pi, d) = R_x(\Pi, d) - R_x(\Pi, \hat{g}^{\Pi}(x))$$

is called the posterior regret of a decision d if a prior is equal to Π.

Definition 1. The predictor $\tilde{g}$ is called the conditional Γ-minimax predictor iff

$$\sup_{\Pi\in\Gamma} R_x(\Pi, \tilde{g}(x)) = \inf_{d\in R}\sup_{\Pi\in\Gamma} R_x(\Pi, d)$$

for every value x of the random variable X.

Definition 2. The predictor $\hat{g}_{PR}$ is called the posterior regret Γ-minimax predictor iff

$$\sup_{\Pi\in\Gamma} U_x(\Pi, \hat{g}_{PR}(x)) = \inf_{d\in R}\sup_{\Pi\in\Gamma} U_x(\Pi, d)$$

for every value x of X.

The definitions are connected with the problem of efficiency of a predictor when a prior runs over a class Γ.

From now on we will suppress x wherever possible in formulas for predictors.

Theorem 1. Let $X = x$. Suppose $\underline{d} = \underline{d}(x) = \inf_{\Pi\in\Gamma} \hat{g}^{\Pi}(x)$ and $\overline{d} = \overline{d}(x) = \sup_{\Pi\in\Gamma} \hat{g}^{\Pi}(x)$ are finite and $\underline{d} < \overline{d}$. Then

$$\hat{g}_{PR} = \underline{d} + \frac{1}{c}\ln\frac{\exp(c(\overline{d} - \underline{d})) - 1}{c(\overline{d} - \underline{d})}$$

and $\hat{g}_{PR} \in (\underline{d}, \overline{d})$.

Proof. For given $X = x$ let $h = h(\Pi, x) = E_{\Pi}(e^{cY}|x)$. Then $\hat{g}^{\Pi} = \frac{1}{c}\ln h$ and

$$U_x(\Pi, d) = he^{-cd} - \ln h + cd - 1.$$

Now proof is similar to the proof of Theorem 1 (Boratyńska 2002).

3. ROBUST PREDICTORS IN POISSON MODEL

Let $X_1, X_2, \ldots, X_n, Y$ be i.i.d. random variables with a Poisson distribution P_θ, where $\theta > 0$ is unknown. Write $X = (X_1, X_2, \ldots, X_n)$. The vector X is observed. A random variable Y is predicted. Let *Gamma*(a, b) denote a Gamma distribution with a density function

$$\pi_{a,b}(\theta) = \frac{b^a}{\Gamma(a)}\theta^{a-1}e^{-b\theta} \quad \text{for} \quad \theta > 0,$$

where $a, b > 0$ are parameters.

Assume that θ has a prior distribution $\Pi_{\alpha_0,\beta_0} = Gamma(\alpha_0, \beta_0)$. If $X = x$ then the posterior distribution is $Gamma(\alpha_0 + T, \beta_0 + n)$, where $T = \sum_{i=1}^{n} x_i$, and the Bayes predictor of the random variable Y under LINEX loss is

$$\hat{g}^{\alpha_0\beta_0} = \frac{1}{c}\ln E_{\Pi_{\alpha_0\beta_0}}(\exp(\theta(e^c - 1))x) =$$

$$= \frac{1}{c}(\alpha_0 + T)\ln\frac{\beta_0 + n}{\beta_0 + n + 1 - e^c}.$$

Note that for a square loss function the value of the Bayes predictor of the random variable Y is equal to the value of the Bayes estimator of the parameter θ. Here, under LINEX loss, the Bayes estimator of θ at a point x is

$$\hat{\theta}^{\alpha_0\beta_0}_{Bay} = \frac{1}{c}(\alpha_0 + T)\ln\frac{\beta_0 + n}{\beta_0 + n - c}.$$

Now suppose that the prior distribution is not exactly specified and consider three classes of priors of θ:

$$\Gamma_1 = \{\Pi_{\alpha_0,\beta} : \Pi_{\alpha_0,\beta} = Gamma(\alpha_0, \beta), \quad \beta \in [\beta_1, \beta_2]\},$$

where $0 < \beta_1 < \beta_2$ are fixed and $\beta_0 \in (\beta_1, \beta_2)$,

$$\Gamma_2 = \{\Pi_{\alpha,\beta_0} : \Pi_{\alpha,\beta_0} = Gamma(\alpha, \beta_0), \quad \alpha \in [\alpha_1, \alpha_2]\},$$

where $0 < \alpha_1 < \alpha_2$ are fixed and $\alpha_0 \in (\alpha_1, \alpha_2)$,

$$\Gamma_3 = \{\Pi_{\alpha,\beta} : \Pi_{\alpha,\beta} = Gamma(\alpha, \beta), \quad \alpha \in [\alpha_1, \alpha_2], \quad \beta \in [\beta_1, \beta_2]\},$$

where $0 < \alpha_1 < \alpha_2$ and $0 < \beta_1 < \beta_2$ are fixed and $\alpha_0 \in (\alpha_1, \alpha_2)$ and $\beta_0 \in (\beta_1, \beta_2)$.

The classes Γ_1, Γ_2 and Γ_3 express three types of uncertainty about the elicited prior.

If a prior is $Gamma(\alpha, \beta)$ then the Bayes predictor of the random variable Y is equal

$$\hat{g}^{\alpha,\beta} = \frac{1}{c}(\alpha + T)\ln\frac{\beta+n}{\beta+n+1-e^c}$$

and it is an increasing function of α and a decreasing function of β. Note that the Bayes predictor exists if $e^c < \beta + n + 1$. Thus from now on we will assume that $c \in (0, \ln(\beta_1 + n + 1))$.

Applying Theorem 1 we obtain the following theorem.

Theorem 2. If the class of priors is equal to Γ_2 then the posterior regret Γ-minimax predictor of Y is equal to

$$\hat{g}_{PR} = \hat{g}^{\alpha_0\beta_0} + \frac{1}{c}\ln\frac{z^{\alpha_2\beta_1}-1}{(\alpha_2-\alpha_1)\ln z},$$

where $z = \dfrac{\beta_0+n}{\beta_0+n+1-e^c}$.

If the class of priors is equal to Γ_1 then the posterior regret Γ-minimax predictor of Y is equal to

$$\hat{g}_{PR} = \hat{g}^{\alpha_0,\beta_0} + \frac{1}{c}\ln\frac{u-1}{\ln u},$$

where $u = \left[\dfrac{(\beta_1+n)(\beta_2+n+1-e^c)}{(\beta_1+n+1-e^c)(\beta_2+n)}\right]^{\alpha_0+T}$, and

$$\hat{g}_{PR} \to \frac{1}{c}\ln\frac{\left(\dfrac{n}{n+1-e^c}\right)^{\alpha^0+T}-1}{(\alpha_0+T)\ln\dfrac{n}{n+1-e^c}} \quad \text{if} \quad \beta_1 \to 0 \quad \text{and} \quad \beta_2 \to \infty.$$

If the class of priors is equal to Γ_3 then the posterior regret Γ-minimax predictor of Y is equal to

$$\hat{g}_{PR} = \hat{g}^{\alpha_1,\beta_2} + \frac{1}{c}\ln\frac{w-1}{\ln w},$$

where $w = \dfrac{(\beta_1+n)^{\alpha_1+T}(\beta_2+n+1-e^c)^{\alpha_2+T}}{(\beta_1+n+1-e^c)^{\alpha_2+T}(\beta_2+n)^{\alpha_1+T}}$.

The next two theorems present the conditional Γ-minimax predictors.

Theorem 3. If the class of priors is equal to Γ_2 then the conditional Γ-minimax predictor of Y is equal to

$$\tilde{g} = \begin{cases} \hat{g}^{\alpha_1,\beta_0} + \frac{1}{c}\ln\frac{(z^{\alpha_2+\alpha_1}-1)(\beta_0+n)}{c(\alpha_2-\alpha_1)} & \text{if} \quad 1-z^{\alpha_1-\alpha_2} \leqslant \frac{c}{\beta_0+n}(\alpha_2-\alpha_1) \\ \hat{g}^{\alpha_2,\beta_0} & \text{otherwise,} \end{cases}$$

where $z = \frac{\beta_0+n}{\beta_0+n+1-e^c}$.

Proof. For a given x a posterior risk of a predictor d is equal to

$$R_z(\Pi_{\alpha,\beta_0}, d) = f(\alpha, d) = e^{-cd}z^{\alpha+T} + cd - c\frac{\alpha+T}{\beta_0+n} - 1,$$

where $z = \frac{\beta_0+n}{\beta_0+n+1-e^c} > 1$. The first and the second derivatives are equal to

$$\frac{\partial f(\alpha,d)}{\partial\alpha} = e^{-cd}z^{\alpha+T}\ln z - \frac{c}{\beta_0+n}$$

and

$$\frac{\partial^2 f(\alpha,d)}{\partial\alpha^2} = e^{-cd}z^{\alpha+T}\ln^2 z.$$

Thus f is a convex function of α and

$$\frac{\partial f(\alpha,d)}{\partial\alpha} = 0 \quad \text{iff} \quad \alpha = \frac{cd + \ln\frac{c}{(\beta_0+n)\ln z}}{\ln z} - T.$$

Hence

$$\sup_{\alpha\in[\alpha_1,\alpha_2]} f(\alpha,d) = \begin{cases} \max(f(\alpha_1,d), f(\alpha_2,d)) & \text{if} \quad d_1 < d < d_2 \\ f(\alpha_2,d) & \text{if} \quad d \leqslant d_1 \\ f(\alpha_1,d) & \text{if} \quad d \geqslant d_2 \end{cases},$$

where $d_i = \hat{g}^{\alpha_i,\beta_0} - \frac{1}{c}\ln\frac{c}{(\beta_0+n)\ln z}$, $i = 1, 2$. The function $l(d) = f(\alpha_1,d) - f(\alpha_2,d)$ is an increasing, continuous function of d and

$$l(d) = 0 \quad \text{iff} \quad d = d^* = \frac{1}{c}\ln\frac{(z^{\alpha_2+T} - z^{\alpha_1+T})(\beta_2+n)}{c(\alpha_2-\alpha_1)},$$

hence

$$\sup_{\alpha\in[\alpha_1\alpha_2]} f(\alpha,d) = \begin{cases} f(\alpha_2,d) & \text{if} \quad d \leqslant d^* \\ f(\alpha_1,d) & \text{if} \quad d \geqslant d^*. \end{cases}$$

If $d < \hat{g}^{\alpha_2,\beta_0}$ *then* $f(\alpha_2, d)$ is a decreasing function of d. Similarly, $f(\alpha_1, d)$ is increasing for $d > \hat{g}^{\alpha_1,\beta_0}$. Thus

$$\inf_{d\in\mathfrak{R}}\sup_{\alpha\in[\alpha_1\alpha_2]} f(\alpha, d) = \begin{cases} f(\alpha_2, d^*), f(\alpha_1, d^*) & \text{if} \quad \hat{g}^{\alpha_1,\beta_0} \leqslant d^* \leqslant \hat{g}^{\alpha_2,\beta_0} \\ f(\alpha_2, \hat{g}^{\alpha_2,\beta_0}) & \text{if} \quad d^* > \hat{g}^{\alpha_2,\beta_0} \\ f(\alpha_1, \hat{g}^{\alpha_1,\beta_0}) & \text{if} \quad d^* < \hat{g}^{\alpha_1,\beta_0}. \end{cases}$$

The inequality

$$d^* \geqslant \hat{g}^{\alpha_1,\beta_0} \tag{1}$$

is equivalent to

$$\frac{z^{\alpha_2-\alpha_1}-1}{\alpha_2-\alpha_1} \geqslant \frac{c}{\beta_0+n}.$$

Consider a function $g(a) = z^{\alpha-\alpha_1}$. Then $h'(a) = z^{\alpha-a_1}$ and there exists $a_0 > \alpha_1$ such that

$$\frac{z^{\alpha_2-\alpha_1}-1}{\alpha_2-\alpha_1} = h'(a_0) > \ln z = \ln\frac{1}{1-\dfrac{e^c-1}{\beta_0+n}} \geqslant \frac{e^c-1}{\beta_0+n} \geqslant \frac{c}{\beta_0+n}.$$

Hence, the inequality (1) is always true.

The inequality $d^* \leqslant \hat{g}^{\alpha_2,\beta_0}$ is equivalent to the inequality

$$1 - z^{\alpha_1-\alpha_2} \leqslant \frac{c}{\beta_0+n}(\alpha_2-\alpha_1),$$

thus we obtain the assertion.

Note that the condition $1 - z^{a_1-\alpha_2} \leqslant \dfrac{c}{\beta_0+n}(\alpha_1-\alpha_1)$ does not depend on observations.

Theorem 4. If the class of priors is equal to Γ_1 then the conditional Γ-minimax predictor of Y is equal to

$$\tilde{g} = \begin{cases} \hat{g}^{\alpha_0,\beta_2} + \dfrac{1}{c}\ln\dfrac{u-1}{r} & \text{if} \quad 1-\dfrac{1}{u} \leqslant r \leqslant u-1 \\ \hat{g}^{\alpha_0,\beta_1} & \text{if} \quad r < 1-\dfrac{1}{u} \\ \hat{g}^{\alpha_0,\beta_2} & \text{if} \quad r > u-1. \end{cases}$$

where $u = \left[\dfrac{(\beta_1+n)(\beta_2+n+1-e^c)}{(\beta_1+n+1-e^c)(\beta_2+n)}\right]^{\alpha_0+T}$ and $r = \dfrac{c(\alpha_0+T)(\beta_2-\beta_1)}{(\beta_1+n)(\beta_2+n)}$.

Proof. Let $b=\beta+n$ and $b_i=\beta_i+n$, $i=1,2$, then b is an increasing function of β and for a given x a Bayes predictor $\hat{g}^{\alpha_0,\beta}$ is a decreasing function of b and

$$R_z(\Pi_{\alpha_0,\beta}, d)=\rho(b,d)=e^{-cd}\left(\frac{b}{b+1-e^c}\right)^{\alpha_0+T}+cd-c\frac{\alpha_0+T}{b}-1\cdot$$

The first derivative is

$$\frac{\partial\rho(b,d)}{\partial b}=\frac{\alpha_0+T}{b^2}\left(e^{-cd}\left(\frac{b}{b+1-e^c}\right)^{\alpha_0+T+1}(1-e^c)+c\right)$$

and

$$\frac{\partial\rho(b,d)}{\partial b}>0\Leftrightarrow\frac{b}{b+1-e^c}<\nu,\quad\text{where}\quad \nu=\left(\frac{e^{cd}c}{e^c-1}\right)^{\frac{1}{\alpha_0+T+1}}.$$

If $\nu-1\leqslant 0$ then ρ is decreasing for $b>0$. If $\nu-1>0$ then ρ has minimum at the point $b_0=\dfrac{\nu(e^c-1)}{\nu-1}$. We have

$$\nu-1>0\Leftrightarrow d>\frac{1}{c}\ln\frac{e^c-1}{c}\cdot$$

Hence

$$\sup_{b\in[b_1,b_2]}\rho(b,d)=\begin{cases}\max(\rho(b_1,d),\rho(b_2,d) & \text{if}\quad d_2<d<d_1\\ \rho(b_2,d) & \text{if}\quad d\geqslant d_1\\ \rho(b_1,d) & \text{if}\quad d\leqslant d_2,\end{cases}$$

where $d_i=\hat{g}^{\alpha_0,\beta_i}+\dfrac{1}{c}\ln\left(\dfrac{e^c-1}{c}\dfrac{b_i}{b_i-1-e^c}\right)$ for $i=1,2$ and $d_2>\dfrac{1}{c}\ln\dfrac{e^c-1}{c}\cdot$

The function $l(d)=\rho(b_1,d)-\rho(b_2,d)$ is decreasing and

$$l(d)=0\Leftrightarrow d=d^*=\hat{g}^{\alpha_0,\beta_2}+\frac{1}{c}\ln\frac{b_1b_2(u-1)}{c(\alpha_0+T)(b_2-b_1)}\cdot$$

Of course $d^*\in[d_2,d_1]$. Hence

$$\sup_{b\in[b_1,b_2]}\pi(b,d)=\begin{cases}\rho(b_1,d) & \text{if}\quad d\leqslant d^*\\ \rho(b_2,d) & \text{if}\quad d\geqslant d^*.\end{cases}$$

Now, analysis similar to that in the proof of Theorem 3 and some calculations give the assertion.

Corollary 1. If the class of priors is equal to Γ_1 and $\beta_1 \to 0$ and $\beta_2 \to \infty$ then the conditional Γ-minimax predictor

$$\tilde{g} \to \begin{cases} \frac{1}{c}\ln\frac{(u_1-1)n}{c(\alpha_0+T)} & \text{if } 1-\frac{1}{u_1} \leqslant \frac{c(\alpha_0+T)}{n} \\ \frac{1}{c}(\alpha_0+T)\ln u_1 & \text{if } \frac{c(\alpha_0+T)}{n} < 1-\frac{1}{u_1}. \end{cases}$$

where $u_1 = \left(\frac{n}{n+1-e^c}\right)^{\alpha_0+T}$.

Now consider the class Γ_3. The derivatives of the posterior risk are as follows

$$\frac{\partial R_x(\Pi_{\alpha,\beta}, d)}{\partial\alpha} = e^{-cd}\left(\frac{\beta+n}{\beta+n+1-e^c}\right)^{\alpha+T}\ln\frac{\beta+n}{\beta+n+1-e^c} - \frac{c}{\beta+n},$$

$$\frac{\partial R_x(\Pi_{\alpha,\beta}, d)}{\partial\beta} = e^{-cd}\frac{(\alpha+T)(1-e^c)}{(\beta+n+1-e^c)^2}\left(\frac{\beta+n}{\beta+n+1-e^c}\right)^{\alpha+T-1} + \frac{c(\alpha+T)}{(\beta+n)^2}$$

and there is no solution of the system of equations

$$\begin{cases} \frac{\partial R_x(\Pi_{\alpha,\beta}, d)}{\partial\alpha} = 0 \\ \frac{\partial R_x(\Pi_{\alpha,\beta}, d)}{\partial\beta} = 0. \end{cases}$$

Thus, for every d supremum of function R_x, when prior runs the class Γ_3, is reached for (α, β) belonging to the boundary of the set $[\alpha_1, \alpha_2]\times[\beta_1, \beta_2]$. Analysis similar to that in the proof of Theorem 3 and 4 shows that

$$\forall d>0 \quad \sup_{\Pi\in\Gamma_3} R_x(\Pi, d) = \max\{R_x(\Pi_{\alpha_i,\beta_j}, d):\quad i=1,2,\quad j=1,2\}.$$

For every $d>0$ we obtain four functions, we choose the maximum and next minimum over d.

Example. Insua et. al. (1999) consider the problem of Bayesian prediction of the number of accidents under the square loss function. They compute the Bayes predictor and its oscillation when a prior runs over a class Γ. The number of accidents is assumed to follow a Poisson model. The parameter θ (representing the accident rate for one policy) is assumed to have a Gamma distribution. The parameters of a prior were approximated using the expert knowledge. As a consequence, a Gamma distribution with parameters $\alpha_0 = 1.59$ and $\beta_0 = 2.22$ is adopted. Since the prior is determined through approximation the family of priors

$$\Gamma = \{Gamma(\alpha, \beta) \colon \ \alpha \in [0.22, 11.1], \ \beta \notin [0.16, 7.95]\}$$

is considered. This family corresponds indeed to wide variations in the parameters. Table 1 presents the data (number of accidents aggregated per year and number of polices).

Table 1

k	Year	Number of polices n_k	Number of accidents T_k
1	1987	4368	75
2	1988	4281	54
3	1989	4157	68
4	1990	3775	60

Using a prior $Gamma(\alpha, \beta)$ and data for $k = 1$, we obtain a posterior distribution $Gamma(\alpha + T_1, \beta + n_1)$, which next becomes a prior for a period $k = 2$ and so on. Thus, in every year we have a new family of priors. Table 2 presents the ends of intervals for parameters α and β in every period.

Table 2

k	Ends of intervals for α		Ends of intervals for β		$\alpha_0 + \sum_{i=1}^{k-1} T_i$	$\beta_0 \sum_{i=1}^{k-1} n_i$
1	0.22	11.1	0.16	7.95	1.59	2.22
2	75.22	86.1	4368.16	4375.95	76.59	4370.22
3	129.22	140.1	8649.16	8656.95	130.59	8651.22
4	197.22	208.1	12806.16	12813.95	198.59	12808.22

Consider the square loss function. If the prior is equal $Gamma(\alpha, \beta)$, and data are T (number of accidents) and n (number of polices) in the period $k-1$, then the Bayes predictor of the number of accidents for n_k polices in the next period is

$$\Lambda_k^{B,K} = n_k \frac{\alpha + T}{\beta + n}.$$

Under LINEX loss function the Bayes predictor is

$$\Lambda_k^{B,K} = n_k \frac{1}{c}(\alpha + T)\ln\frac{\beta + n}{\beta + n + 1 - e^c}.$$

Note that in our consideration above the example, we predict the number of accidents (claims) for one policy. Thus here we must multiply every predictor by the number of polices.

From now on let $\Lambda_k^{B,K}$ denote the Bayes predictor of number of accidents in the period k under the square loss function and $\Lambda_k^{B,K}$ the Bayes predictor under LINEX loss function, both if the prior is $Gamma(\alpha_0 + \sum_{i=1}^{k-1} T_i, \beta_0 + \sum_{i=1}^{k-1} n_i)$. Let $\Lambda_{PR,k}$ and $\tilde{\Lambda}_k$ denote the posterior regret Γ-minimax predictor and the conditional Γ-minimax predictor under LINEX loss for n_k polices.

Table 3

Square loss function					
k	T_k	$\Lambda_k^{B,K}$	Bounds of predictor		Oscillation
1	75				
2	54	75.0	73.6	84.4	10.8
3	68	62.7	62.1	67.3	5.2
4	60	58.5	58.1	61.3	3.2
LINEX loss $c = 0.001$					
k	T_k	$\Lambda_k^{B,L}$	Bounds of predictor		Oscillation
1	75				
2	54	75.1	73.6	84.4	10.8
3	68	62.8	62.1	67.4	5.3
4	60	58.6	58.1	61.4	3.3
LINEX loss $c = 0.01$					
k	T_k	$\Lambda_k^{B,L}$	Bounds of predictor		Oscillation
1	75				
2	54	75.4	73.9	84.8	10.9
3	68	63.1	62.4	67.7	5.3
4	60	58.8	58.4	61.7	3.3
LINEX loss $c = 0.1$					
k	T_k	$\Lambda_k^{B,L}$	Bounds of predictor		Oscillation
1	75				
2	54	78.9	77.4	88.7	11.3
3	68	66.0	65.3	70.8	5.5
4	60	61.6	61.1	64.5	3.4

Table 3 presents the values of Bayes predictors $\Lambda_k^{B,K}$ and $\Lambda_k^{B,L}$ for some values of c and oscillation of Bayes predictors when a prior runs over the class of priors (see Table 2). The difference between values of predictor under LINEX loss and under the square loss is an increasing function of c, for $c > 0$. The oscillation decreases if k increases.

Table 4 presents the values of the posterior regret Γ-minimax predictor and the conditional Γ-minimax predictor. For our data the conditional Γ-minimax predictor is equal to the Bayes predictor corresponding to the prior with parameters $\alpha_2 = 11.1$ and $\beta_1 = 0.16$.

Table 4

k	c = 0.001		c = 0.01		c = 0.1	
	$\Lambda_{PR,k}$	$\tilde{\Lambda}_k$	$\Lambda_{PR,k}$	$\tilde{\Lambda}_k$	$\Lambda_{PR,k}$	$\tilde{\Lambda}_k$
2	79.0	84.4	79.4	84.8	83.1	88.7
3	64.7	67.4	65.0	67.7	68.0	70.8
4	59.8	61.4	60.0	61.7	62.8	64.5

Now consider the collective risk model. We have a sequence of random variables: $Y, Z_1, Z_2, \ldots$ A random variable Y describes the number of claims of a given contract or a portfolio of contracts in a future period. Random variables Z_i, $i = 1, 2, 3, \ldots$ describe sizes of claims. Random variables $Y, Z_1, Z_2, \ldots$ are independent and Z_i, $i = 1, 2, 3, \ldots$ are identically distributed. Assume we know the probability distribution function of Z_i. Let $S = \sum_{i=1}^{Y} Z_i$. Assume that S is to be predicted by a function G depending on $X = (X_1, X_2, \ldots, X_n)$, where $X_1, X_2, \ldots, X_n$ are numbers of claims in previous periods. Again use the LINEX function as a criterion function. Obtained predictor can be interpreted as amount of premium.

Assume X and Y satisfy the conditions presented in the beginning of the Section 3. If $X = x$ and Π is a prior distribution of θ then the Bayes predictor of S is equal to

$$G^{\Pi} = \frac{1}{c} \ln E_{\Pi}(e^{cS}|x) = \frac{1}{c} \ln E_{\Pi}(e^{Y \ln M_z(c)}|x),$$

where $M_z(c)$ denotes the moment generation function of a random variable Z at the point c. Assume that $M_z(c)$ exists and it is finite. For $\Pi = Gamma(\alpha, \beta)$ we obtain

$$G^{\alpha,\beta} = \frac{1}{c}(\alpha + T)\ln\frac{\beta + n}{\beta + n + 1 - M_z(c)}.$$

Assume that c satisfies: $M_z(c) < \beta_1 + n + 1$ and $M_z(c) > 1$. Then we obtain the following corollaries. The proofs of corollaries are like proofs of Theorems 2, 3 and 4. We only remember that if $\Pi = Gamma(\alpha, \beta)$ then

$$E_\Pi(e^{cS}|x) = \left(\frac{\beta + n}{\beta + n + 1 - M_z(c)}\right)^{\alpha + T}$$

and

$$E_\Pi(S|x) = EZ\frac{\alpha + T}{\beta + n}.$$

Corollary 2. If the class of priors is equal to Γ_2 then the posterior regret Γ-minimax predictor of S is equal to

$$G_{PR} = G^{\alpha_1\beta_0} + \frac{1}{c}\ln\frac{z^{\alpha_2-\alpha_1} - 1}{(\alpha_2 - \alpha_1)\ln z},$$

where $z = \dfrac{\beta_0 + n}{\beta_0 + n + 1 - M_z(c)}$.

If the class of priors is equal to Γ_1 then the posterior regret Γ-minimax predictor of S is equal to

$$G_{PR} = G^{\alpha_0\beta_2} + \frac{1}{c}\ln\frac{u - 1}{\ln u},$$

where

$$u = \left[\frac{(\beta_1 + n)(\beta_2 + n + 1 - M_z(c))}{(\beta_1 + n + 1 - M_z(c))(\beta_2 + n)}\right]^{\alpha_0 + T}.$$

If the class of priors is equal to Γ_3 then the posterior regret Γ-minimax predictor of S is equal to

$$G_{PR} = G^{\alpha_1\beta_2} + \frac{1}{c}\ln\frac{w - 1}{\ln w},$$

where

$$w=\left[\frac{\beta_1+n}{\beta_1+n+1-M_z(c)}\right]^{\alpha_2+T}\left[\frac{\beta_2+n+1-M_z(c)}{\beta_2+n}\right]^{\alpha_1+T}.$$

Corollary 3. If the class of priors is equal to Γ_2 then the conditional Γ-minimax predictor of S is equal to

$$\tilde{G}=\begin{cases} G^{\alpha_1,\beta_0}+\frac{1}{c}\ln\frac{(z-1)(\beta_0+n)}{cEZ(\alpha_2-\alpha_1)} & \text{if} \quad \frac{1-z^{-1}}{\alpha_2-\alpha_1}\leqslant\frac{cEZ}{\beta_0+n}\leqslant\frac{z-1}{\alpha_2-\alpha_1} \\ G^{\alpha_1,\beta_0} & \text{if} \quad \frac{1-z^{-1}}{\alpha_2-\alpha_1}<\frac{cEZ}{\beta_0+n} \\ G^{\alpha_2,\beta_0} & \text{otherwise} \end{cases},$$

where $z=\left[\frac{\beta_0+n}{\beta_2+n+1-M_z(c)}\right]^{\alpha_2-\alpha_1}$.

If the class of priors is equal to Γ_1 then the conditional Γ-minimax predictor of S is equal to

$$\tilde{G}=\begin{cases} G^{\alpha_0,\beta_2}+\frac{1}{c}\ln\frac{u-1}{r} & \text{if} \quad 1-\frac{1}{u}\leqslant r\leqslant u-1 \\ G^{\alpha_0,\beta_1} & \text{if} \quad r<1-\frac{1}{u} \\ G^{\alpha_0,\beta_2} & \text{if} \quad r>u-1 \end{cases},$$

where $u=\left[\frac{(\beta_1+n)(\beta_2+n+1-M_z(c))}{(\beta_1+n+1-M_z(c))(\beta_2+n)}\right]^{\alpha_0+T}$ and $r=\frac{cEZ(\alpha_0+T)(\beta_2-\beta_1)}{(\beta_1+n)(\beta_2+n)}$.

Note that for every fixed x there exist priors belonging to the considered class Γ such that the posterior regret Γ-minimax predictor and the conditional Γ-minimax predictor are equal to the Bayes predictors with respect to those priors.

REFERENCES

Berger J. O. (1990), "Robust Bayesian Analysis: Sensitivity to the Prior", *Journal of Statistical Planning and Inference*, **25**, 303–328.

Berger J. O. (1994), "An Overview of Robust Bayesian Analysis", *Test*, **3**, 5–124 (with discussion).

Boratyńska A. (2002), "Posterior Regret Γ-minimax Estimation in a Normal Model with Asymmetric Loss Function", *Applicationes Mathematicae*, **29**, 7–13.

Gómez-Déniz E., Hernández-Bastida A., Vázquez-Polo F. J. (1999), "The Esscher Premium Principle in Risk Theory: A Bayesian Sensitivity Study", *Insurance: Mathematics and Economics*, **25**, 387–395.

Gómez-Déniz E., Hernández-Bastida A., Vázquez-Polo F. J. (2002), "Bounds for Ratios of Posterior Expectations: Applications in the Collective Risk Model", *Scandinavian Actuarial Journal*, 37–44.

Goovaerts M. J. (1990), *Effective Actuarial Methods*, North Holland, Amsterdam.

Insua R. S., Martin J., Insua R. D., Ruggeri F. (1999), "Bayesian Forecasting for Accident Proneness Evaluation", *Scandinavian Actuarial Journal*, 134–156.

Insua R. D., Ruggeri F. (eds), (2000), "Robust Bayesian Analysis", *Lectures Notes in Statistics*, **152**, Springer-Verlag, New York.

Klugman S. A. (1992), *Bayesian Statistics in Actuarial Science, with Emphasis on Credibility*, Kluwer Academic Publishers.

Klugman S. A., Panjer H., Willmot G. (1998), *Loss Models from Data to Decisions*, John Wiley and Sons, New York.

Makov U. E., Smith A. F. M., Liu Y-H. (1996), "Bayesian Methods in Actuarial Science", *The Statistician*, **45**, 503–515.

Zellner A. (1986), "Bayesian Estimation and Prediction Using Asymmetric Loss Functions", *Journal American Statistical Association*, **81**, 446–451.

Agata Boratyńska

ODPORNA PREDYKCJA BAYESOWSKA PRZY ASYMETRYCZNEJ FUNKCJI STRATY W MODELU POISSONA DLA RYZYKA UBEZPIECZENIOWEGO

(Streszczenie)

W odpornej analizie bayesowskiej rozkład *a priori* nie jest dokładnie wyznaczony, ale należy do pewnej rodziny Γ rozkładów *a priori*. Przy takim założeniu otrzymujemy również rodzinę decyzji bayesowskich. Celem jest natomiast wybór jednej reguły „optymalnej".

W artykule rozważany jest problem odpornej predykcji bayesowskiej zmiennej losowej o rozkładzie Poissona przy funkcji straty LINEX. Niedokładność w wyznaczeniu rozkładu *a priori* modeluje się za pomocą trzech rodzin rozkładów *a priori*. Wyznaczamy predyktor warunkowo Γ-minimaksowy i predyktor o Γ-minimaksowej utracie *a posteriori*. Podajemy zastosowania w kolektywnym modelu ryzyka.