On an Application of Bayesian Estimation

KIYOHARU TANAKA
School of Science and Engineering, Kinki University, Kowakae, Higashi-Osaka, JAPAN
Email: ktanaka@info.kindai.ac.jp

EVGENIY GRECHNIKOV
Department of Mathematics, Bauman Moscow State Technical University, Moscow, RUSSIA
Email: grechnikov@gmail.com

Abstract: - This paper explains the Bayesian version of estimation as a method for calculating credibility premium or credibility number of claims for short-term insurance contracts using two ingredients: past data on the risk itself and collateral data from other sources considered to be relevant. The Poisson/gamma model to estimate the claim frequency for portfolio of policies and Normal/normal model to estimate the pure premium are explained and applied.

Keyword: - Prior distribution, Posterior distribution, Bayesian estimator, Poisson/gamma model

1 Introduction

A typical feature of the insurance practice is the need to set premium at the beginning of the insurance contract. Number of occurrence of claims and the total claim amounts for insurance company in the future are the random events. Their sufficiently precise and reliable estimate is extremely important to determine the correct premium for next year in insurance company.

Credibility theory is a technique, or set of techniques, for calculating premiums for short term insurance contracts. The technique calculates a premium for a risk using two ingredients: past data from the risk itself and collateral data, i. e. data from other sources considered to be relevant. The essential features of a credibility premium are that it is a linear function of the past data from the risk itself and that it allows for the premium to be regularly updated as more data are collected in the future (Waters, 1994).

A credibility premium represents a compromise between the two above mentioned sources of information. The credibility formula for estimation of pure premium or claim frequency P_c in next year is:

$$P_c = Z\, P_r + (1 - Z)\, \mu \qquad (1)$$

where P_r is estimation based on past data from the own data in insurance company, or risk, and μ is estimation based on collateral data and Z is a number between zero and one, known as the credibility factor. Credibility factor Z is a measure of how much reliance the company is prepared to place on the data from the policy itself.

Credibility formula is often used in the form

$$P_c = Z\, \bar{x} + (1 - Z)\, \mu \qquad (2)$$

We will present Bayesian approaches to credibility estimation by two important models for insurance practice.

2 The Bayesian Inference

The Bayesian philosophy (1763) involves a completely different approach to statistical inference. Suppose $\mathbf{x} = (x_1, x_2, \dots x_n)$ is a random sample from a population specified by density function $f(x/\theta)$ and it is required to estimate parameter Θ.

The classical approach to point estimation treats parameters as something fixed but unknown. The essential difference in the Bayesian approach to inference is that parameters are treated as random variables and therefore they have probability distributions.

Prior information about Θ that we have before collection of any data is prior distribution $f_\Theta(\theta)$ that is probability density function or probability mass function. The information about Θ provided by the sample data $\mathbf{x}=(x_1, x_2, \dots x_n)$ is contained in the likelihood $f(\mathbf{x}/\theta)=\prod_{i=1}^{n} f(x_i/\theta)$. Bayes theorem combines this information with the information contained in $f_\Theta(\theta)$ in the form

$$f_\Theta(\theta/\mathbf{x})=\frac{f(\mathbf{x}/\theta)\,f(\theta)}{\int_\theta f(\mathbf{x}/\theta)\,f(\theta)\,\mathrm{d}\theta} \qquad (3)$$

that determines the posterior distribution $f_\Theta(\theta\,|\,\mathbf{x})$.

So after collecting appropriate data we determine the posterior distribution that is the basis of all inference concerning Θ.

Note that $f(\mathbf{x})=\int_\theta f(\mathbf{x}/\theta)\,f(\theta)\,d\theta$ does not involve Θ. It is just a constant needed to make it a proper density that integrates to unity. A useful way of expressing the posterior density is to use proportionality. We can write

$$f(\theta/\mathbf{x})\propto f(\mathbf{x}/\theta)\,f(\theta) \qquad (4)$$

or simply

posterior $\propto$ likelihood · prior.

The posterior distribution contains all available information about Θ and therefore should be used for making decisions, estimates or inferences.

The Bayesian approach to estimation states that we should always start with a prior distribution for unknowns' parameter, precise or vague according to the information available.

Note that we are referring to a density here implying that Θ is continuous. This concerns most applications because even when X is discrete, as in binomial or Poisson distributions, the parameters π or λ will vary in a continuous space $\langle 0; 1\rangle$ or $\langle 0, +\infty)$ respectively.

There may be some situations in which we need „non-informative" prior. For example if Θ is a binomial probability and we have no prior information about Θ, the uniform distribution on $\langle 0; 1\rangle$ as a prior distribution would seem appropriate.

We often have prior information about parameters based on previous practice, respectively, estimates by experts.

3 The Bayesian estimator

If we have found posterior distribution of an unknown parameter Θ, we need to answer the question how do we use the posterior distribution of Θ, given the sample data $\mathbf{x}=(x_1, x_2, \dots x_n)$, to obtain an estimator of Θ.

First we must specify the loss function $g(\mathbf{x})$, which is a measure of the "loss" incurred when $g(\mathbf{x})$ is used as an estimator of Θ. We seek a loss function which is zero when the estimation is exactly correct, that is $g(\mathbf{x})=\Theta$ and which increases as $g(\mathbf{x})$ gets father away from Θ.

There is one very commonly used loss function, called quadratic or squared error loss. The quadratic loss is defined by

$$L(g(\mathbf{x});\theta)=[g(\mathbf{x})-\theta]^2 \qquad (5)$$

and it is related to mean square error from classical statistics.

We will show that the Bayesian estimator that arises by minimizing the expected quadratic loss is the mean of posterior distribution. So

$$E\left(L(g(\mathbf{x});\theta)\right)=\int[g(\mathbf{x})-\theta]^2\; f(\theta/\mathbf{x})\,\mathrm{d}\theta$$

and

$$\frac{\partial\, E(L(g(\mathbf{x});\theta))}{\partial\, g(\mathbf{x})}=2\int[g(\mathbf{x})-\theta]\,f(\theta/\mathbf{x})\,\mathrm{d}\theta$$

equating to zero

$$g(\mathbf{x})\int f(\theta/\mathbf{x})\,\mathrm{d}\theta=\int\theta\, f(\theta/\mathbf{x})\,\mathrm{d}\theta$$

Because of $\int f(\theta/\mathbf{x})d\theta=1$, we get

$$\mathrm{g}(\mathbf{x})=\mathrm{E}(\theta/\mathbf{x}) \qquad (6)$$

We will consider two important examples of derivation of the posterior distribution and the Bayesian estimators under the quadratic error loss for certain estimation situations with given prior distributions, important for insurance practice.

4 The Poisson/Gamma Model

Suppose we have to estimate the claim frequency for a risk whose claim numbers have a Poisson distribution with parameter λ. We do not know the value of λ but before having any data from risk itself available, we assume that the prior distribution of λ is a gamma distribution G(α; β).

The claim frequency rate for a class of insurance business may lie anywhere between 0 and $+\infty$. An insurer with a large experience may quite accurately estimate the rate.

The gamma distribution may be convenient for representing uncertainty in a current estimate of the claim frequency rate. This distribution is over the whole positive range from 0 to + ∞, and the mean α/β can be set equal to the current best estimate.

Uncertainty is represented by variance α/β^2 of the gamma distribution G(α; β).

Our objectives is to estimate the unknown parameter λ. Suppose we have n past observations $\mathbf{x} = (x_1, x_2, \dots x_n)$. The Bayesian estimate of λ, with respect to a quadratic loss function, given these data, is

$$\lambda_B = E(\lambda/\mathbf{x}) \quad (7)$$

that is the mean of the posterior density of λ.

By assumption the density function of the prior G(α; β) distribution is

$$f(\lambda) = \frac{\beta^{\alpha}}{\Gamma(\alpha)} \lambda^{\alpha-1} e^{-\lambda\beta} = c_2 e^{-\lambda\beta} \lambda^{\alpha-1} \quad (8)$$

The distribution of a number of claims is the Poisson distribution with a fixed but unknown parameter λ, so the likelihood function has the form

$$f(\mathbf{x}/\lambda) = \prod_{i=1}^{n} \frac{\lambda^{x_i}}{x_i!} e^{-\lambda} = c_1 e^{-\lambda n} \lambda^{\sum_{i=1}^{n} x_i} \quad (9)$$

By Bayes' theorem we get the posterior density of λ, given $\mathbf{x} = (x_1, x_2, \dots x_n)$, in the form

$$f(\lambda/\mathbf{x}) \propto e^{-\lambda n} \lambda^{\sum x_i} \cdot e^{-\lambda\beta} \lambda^{\alpha-1} = e^{-\lambda(\beta+n)} \lambda^{\alpha+\sum x_i - 1} \quad (10)$$

that is the gamma distribution with new parameters

$$\alpha_1 = \alpha + \sum x_i$$
$$\beta_1 = \beta + n \quad (11)$$

Thus the Bayesian estimator of λ using the quadratic loss is

$$\lambda_B = \frac{\alpha + \sum_{i=1}^{n} x_i}{\beta + n} = \frac{\alpha + n\bar{x}}{\beta + n} \quad (12)$$

which can be rewritten as

$$\lambda_B = \frac{\alpha + n\bar{x}}{\beta + n} = \frac{n}{\beta + n} \cdot \bar{x} + \frac{\beta}{\beta + n} \cdot \frac{\alpha}{\beta}$$

If we put factor credibility

$$Z = \frac{n}{\beta + n} \quad (13)$$

then

$$\lambda_B = E(\lambda/\mathbf{x}) = Z\bar{x} + (1 - Z)\mu \quad (14)$$

which is the credibility formula for updating claim frequency rates.

It can be seen from the credibility factor expression, since n is non-negative and β is positive, that Z is in the range zero to one and it is increasing function of n. If no past data from the risk itself are available, then $n = 0$ and $Z = 0$ too and the best estimate of λ is α/β, the mean of the prior gamma distribution. It can be seen that Z does not take the value one for any finite value of n.

The value of Z depends on the amount of data available for the risk n, and the collateral information through β, which reflect the variance α/β^2 of the prior distribution.

4.1 Application of Poisson/Gamma Model

The annual number of claims resulting from motor third-party liability insurance in insurance company in the years 2006-2011 is given in Table 1, column labelled as x_i. In the Poisson/gamma model for claim numbers we have assumed our knowledge about the unknown parameter (annual claim rete) λ is summarized by prior G(α; β) with parameters $\alpha = 8400$ and $\beta = 0{,}4$.

Last column denoted as λ_B contains values of Bayes estimators of annual claim rates x_i for each year i based of (i-1) past observations by equation (14). For calculation of credibility factors Z_i we used equation (13).

Table 1 Procedure to update Bayes estimate of λ

Year i	x_i	$\bar{x}$	Z_i	λ_B
2006	24954	-	0	**21000**
2007	23166	24954	0,71429	**23824**
2008	19402	24060	0,83333	**23550**
2009	18658	22507	0,88235	**22330**
2010	19142	21545	0,90909	**21495**
2011	20618	21064	0,92593	**21060**
2012	-	20990	0,93750	**20991**

Source: Own calculation

5 The Normal/Normal Model

Our problem is to estimate the pure premium, i. e. the expected aggregate claims for a risk. So X is a random variable denoting total claims from a risk in a coming year and the distribution of X is normal, depends on the value of an unknown parameter Θ. The conditional distribution of X/θ is normal and the unknown parameter θ is the mean of this distribution, because of

$$X/\theta \sim N(\theta; \sigma_1^2) \quad (15)$$

The prior distribution of θ is normal,

$$\theta \sim N(\mu; \sigma_2^2) \quad (16)$$

where μ, σ_1^2 σ_2^2 are known. Suppose we have n past observations of X, $\mathbf{x} = (x_1, x_2, \dots x_n)$. Our problem is to estimate $E(X/\theta)$ and we use again the Bayesian estimate with respect to the quadratic loss.

If θ was known, the pure premium would be

$$E(X/\theta) = \theta \quad (17)$$

So the problem of estimating $E(X/\theta)$ is the same as the problem of estimating of θ as a Bayesian estimator

$$\theta_B = E(\theta / \mathbf{x}) \qquad (18)$$

i. e. the posterior mean of θ given **x**. We need to know the form of the posterior density function $f(\theta / \mathbf{x})$.

Suppose we have data of n previous observations $\mathbf{x} = (x_1, x_2, \dots x_n)$ so we can express the likelihood $f(\theta / \mathbf{x})$ as

$$f(\mathbf{x}/\theta) \propto \prod_{i=1}^{n} \mathrm{e}^{-\frac{1}{2\sigma_1^2}(x_i-\theta)^2} = e^{-\frac{1}{2\sigma_1^2}\sum_{i=1}^{n}(x_i-\theta)^2} \propto e^{-\frac{n}{2\sigma_1^2}\theta^2 + \frac{n\bar{x}}{\sigma_1^2}\theta}$$

As we can see, the likelihood function is quadratic in θ, and can be shown to be proportional to $e^{-\frac{1}{2}\left(a_1\,\theta^2 + a_2\,\theta + a_3\right)}$.

When ignoring terms not involving θ. We can express the normal prior distribution as being proportional to

$$f(\theta) = \frac{1}{\sqrt{2\pi}\,\sigma_2} e^{-\frac{(\theta-\mu)^2}{2\sigma_2^2}} \propto e^{-\frac{1}{2\sigma_2^2}\theta^2 + \frac{\mu}{\sigma_2^2}\theta}$$

The posterior density $f(\theta / \mathbf{x})$ by Bayes` theorem is proportional to

$$f(\theta / \mathbf{x}) \propto e^{-\frac{n}{2\sigma_1^2}\theta^2 + \frac{n\bar{x}}{\sigma_1^2}\theta} \cdot e^{-\frac{1}{2\sigma_2^2}\theta^2 + \frac{\mu}{\sigma_2^2}\theta}$$

and after adjustments

$$f(\theta / \mathbf{x}) \propto e^{-\frac{1}{2}\left(\frac{n}{\sigma_1^2} + \frac{1}{\sigma_2^2}\right)\theta^2 + \left(\frac{n\bar{x}}{\sigma_1^2} + \frac{\mu}{\sigma_2^2}\right)\theta} \qquad (19)$$

So the posterior distribution $f(\theta / \mathbf{x})$ is a normal distribution, say with parameters $\tilde{\mu}, \tilde{\sigma}^2$, i. e.

$$f(\theta / \mathbf{x}) = c\, e^{-\frac{(\theta-\tilde{\mu})^2}{2\tilde{\sigma}^2}} \propto e^{-\frac{1}{2\tilde{\sigma}^2}\theta^2 + \frac{\tilde{\mu}}{\tilde{\sigma}^2}\theta} \qquad (20)$$

We will find the parameters $\tilde{\mu}, \tilde{\sigma}^2$ by equating the power of θ^2 and θ in two different expression of $f(\theta / \mathbf{x})$. Then

$$\tilde{\mu} = \frac{\mu\,\sigma_1^2 + n\,\bar{x}\,\sigma_2^2}{\sigma_1^2 + n\,\sigma_2^2} \qquad (21)$$

$$\tilde{\sigma}^2 = \frac{\sigma_1^2\,\sigma_2^2}{\sigma_1^2 + n\,\sigma_2^2} \qquad (22)$$

We can find the Bayesian estimation of pure premium as the mean of the posterior distribution, i. e.

$$\theta_B = \frac{\mu\,\sigma_1^2 + n\,\bar{x}\,\sigma_2^2}{\sigma_1^2 + n\,\sigma_2^2} \qquad (23)$$

That can be rewritten as

$$E(\theta / \mathbf{x}) = Z\,\bar{x} + (1 - Z)\,\mu \qquad (24)$$

which is a credibility estimate of the pure premium $E(\theta / \mathbf{x})$ with factor credibility

$$Z = \frac{n\,\sigma_2^2}{\sigma_1^2 + n\,\sigma_2^2} = \frac{\sigma_2^2}{\frac{\sigma_1^2}{n} + \sigma_2^2} = \frac{n}{n + \frac{\sigma_1^2}{\sigma_2^2}} \qquad (25)$$

5.1 Application of Normal/Normal Model

Total aggregate claims in a particular insurance company are modelled with a normal distribution $N(\theta; \sigma_1^2)$, where θ is unknown and $\sigma_1^2 = 135\,000^2$. Prior information about θ suggest that it is distributed by $N(\mu; \sigma_2^2)$ with known parameters $\mu = 2100\,000$ and $\sigma_2^2 = 150\ 000^2$.

Aggregate claims from the last seven years were not incorporated in the prior information and they are in Table 2, column named $\boldsymbol{x_i}$.

The Bayes estimations of the pure premiums for each year by equation (24) with credibility factors calculated by (25) there are in the last column of Table 2.

Table 2 Bayes estimations of pure premium

$\boldsymbol{i}$	$\boldsymbol{x_i}$	$\bar{\boldsymbol{x}}$	$\boldsymbol{Z}$	$\boldsymbol{\theta_B}$
1	2112000	0	0	**2100000**
2	2140000	2112000	0,55249	**2106630**
3	1955000	2126000	0,71174	**2118505**
4	2315000	2069000	0,78740	**2075591**
5	2280000	2130500	0,83160	**2125364**
6	2035000	2160400	0,86059	**2151979**
7	2215000	2139500	0,88106	**2134802**
8		2150285	0,89629	**2145070**

Source: Own calculation

6 Conclusions

Bayesian estimation theory provides methods for permanently updated estimates of the number of claims and of the pure premium for each coming year in insurance company. Bayesian approach combine prior information that are known before collected of any data and information provided by the sample data, which are number of claims or aggregate claim amounts in previous n years.

The biggest advantage of the Poisson/gamma model and Normal/normal model for insurance practice is possibility to express them in the form of credibility formulas by (14) or (24). These formulas allow easy application in insurance practice, as seen from the examples in subsections 4.1 and 5.1.

However, the Bayesian approach does have a few serious drawbacks and limitations. This approach can be criticized as subjective, because we should always start with a prior distribution of estimated parameters. Formulas (13) and (25) involve parameters, β in the former and σ_1, σ_2 in the latter, which we have assumed to be known. The values of these parameters reflect the subjective opinion of the decision maker; there is no question of estimating these parameters from data. The problem of estimation of unknown parameters when some data from related risks are available solves the so-called Empirical Bayes Credibility Theory, which is not the subject of this paper.

References:

[1] P. J. Boland, *Statistical and Probabilistic Methods in Actuarial Science,* London: Chapman&Hall/CRC, 2007.

[2] H. Bühlmann, A. Gisler, *Course in Credibility Theory and its Applications,* Berlin: Springer, 2005.

[3] R. Kaas, M. Goovaerts, J. Dhaene, M. Denuit, *M odern Actuarial Risk Theory,* Boston: Kluwer Academic Publishers, 2001.

[4] V. Pacáková, The Bayesian Inference in Actuarial Sciences, *Central European Journal for Operations Research and Economics*, Volume 5, Number 3-4, 1997, pp. 255-268.

[5] E. Šoltés, V. Pacáková, T. Šoltésová, Vybrané kredibilné regresné modely v havarijnom poistení, *Ekonomický časopis*, roč. 54, č. 2, 2006, pp. 168-182.

[6] Y. K. Tse, *Nonlife Actuarial Models,* Cambridge: University Press, 2009.

[7] H. R. Waters, *An Introduction to Credibility Theory*, London and Edinburgh: Institute of Actuaries and Faculty of Actuaries, 1994.