Lecture

INSTITUTE OF ACTUARIAL & QUANTITATIVE STUDIES

Class: MSc

Subject : Statistical and Risk Modelling - 2

Chapter: Unit 2 Chapter 1

Chapter Name: Applying Bayesian Modelling and Credibility Theory

Topics to be covered

1. Introduction
2. Bayesian Estimate
3. Bayesian Inference & Estimation
4. Posterior Distribution of Parameter
5. Bernoulli Random Variable

1 Introduction

- Bayesian statistics is an approach to data analysis based on Bayes' theorem, where available knowledge about parameters in a statistical model is updated with the information in observed data.
- The background knowledge is expressed as a prior distribution and combined with observational data in the form of a likelihood function to determine the posterior distribution. The posterior can also be used for making predictions about future events.
- This Primer describes the stages involved in Bayesian analysis, from specifying the prior and data models to deriving inference, model checking and refinement.

2 Bayesian Estimate

- The Bayesian estimate of the risk parameter under the squared-error loss function is the mean of the posterior distribution.
- Likewise, the Bayesian estimate of the mean of the random loss is the posterior mean of the loss conditional on the data.
- In general, the Bayesian estimates are difficult to compute, as the posterior distribution may be quite complicated and intractable. There are, however, situations where the computation may be straightforward, as in the case of conjugate distributions.

3 Bayesian Inference and Estimation

- The classical and Bühlmann credibility models update the prediction for future losses based on recent claim experience and existing prior information.
- In these models, the random loss variable X has a distribution that varies with different risk groups.
- Based on a sample of n observations of random losses, the predicted value of the loss for the next period is updated.
- The predictor is a weighted average of the sample mean of X and the prior mean, where the weights depend on the distribution of X across different risk groups.

3 Bayesian Inference and Estimation

- We formulate the aforementioned as a statistical problem suitable for the Bayesian approach of statistical inference and estimation. The set-up is summarized as follows:

    i. Let X denote the random loss variable (such as claim frequency, claim severity, and aggregate loss) of a risk group. The distribution of X is dependent on a parameter θ, which varies with different risk groups and is hence treated as the realization of a random variable Θ.
    ii. Θhas a statistical distribution called the prior distribution. The prior pdf of Θ is denoted by $f_{\Theta}(\theta|\gamma)$ (or simply f_{Θ}(θ)), which depends on the parameter γ , called the hyperparameter.
    iii. The conditional pdf of X given the parameter θ is denoted by $f_{(X|\Theta)}$(x | θ). Suppose X = $\{X_1, \dots X_n\}$ is a random sample of X , and x = $(x_1, .. x_n)$ is a realization of X. The conditional pdf of X is :

$$f_{(X|\Theta)}(x|\theta) = \prod_{i=1}^{n} f_{X|\Theta}(x_i|\theta)$$

We call $f_{x|\Theta}(x \mid \theta)$ the likelihood function.

iv. iv. Based on the sample data x, the distribution of Θ is updated. The conditional pdf of ▯ given x is called the posterior pdf, and is denoted by $f_{\Theta|X}\,(\theta \mid x)$.
v. 5 An estimate of the mean of the random loss, which is a function of Θ, is computed using the posterior pdf of Θ. This estimate, called the Bayes estimate, is also the predictor of future losses.

3 Bayesian Inference and Estimation

- Bayesian inference differs from classical statistical inference in its treatment of the prior distribution of the parameter θ. Under classical statistical inference, θ is assumed to be fixed and unknown, and the relevant entity for inference is the likelihood function.
- For Bayesian inference, the prior distribution has an important role. The likelihood function and the prior pdf jointly determine the posterior pdf, which is then used for statistical inference.

4 Posterior distribution of parameter

- Given the prior pdf of Θ and the likelihood function of X, the joint pdf of Θ and X can be obtained as follows
$$f_{\Theta X}(\theta, x) = f_{(X|\Theta)}(x \mid \theta) * f\_\Theta\,(\theta)$$
- Integrating out θ from the joint pdf of Θ and X, we obtain the marginal pdf of X as
$$f_X(x) = \int_{\theta \epsilon \Omega_\Theta} f_{(X|\Theta)}(x|\theta)\, f_\Theta(\theta)$$
- where Ω_Θ is the support of Θ.
- Now we can turn the question around and consider the conditional pdf of Θ given the data x, i.e. f□| X (θ | x). Combining the above equations , we have
$$f_{\Theta|X}(\theta|x) = \frac{f_{\Theta X}(\theta, x)}{f_X(x)}$$
$$= \frac{f_{X|\Theta}(x|\theta) f_\Theta(\theta)}{\int_{\theta \epsilon \Omega_\Theta} f_{X|\Theta}(x|\theta)\, f_\Theta(\theta) d\theta}$$
- The posterior pdf describes the distribution of Θ based on prior information about Θ and the sample data x. Bayesian inference about the population as described by the risk parameter Θ is then based on the posterior pdf.

5 Bernoulli Random Variable

- Example:

  Let X be the Bernoulli random variable which takes value 1 with probability θ and 0 with probability 1 − θ. If Θ follows the beta distribution with parameters α and β, i.e. Θ ∼ B(α, β), calculate the posterior pdf of Θ given *X*.

- Solution:

  As X is Bernoulli, the likelihood function of X is

  $f_{X|\Theta}(x \mid \theta) = \theta^{x}(1-\theta)^{1-x}$, for x = 0, 1.

  Since Θ is assumed to follow the beta distribution with hyperparameters α and β, the prior pdf of Θ is

  $f_{\Theta}(\theta) = \frac{\theta^{\alpha-1}(1-\theta)^{\beta-1}}{B(\alpha,\beta)}$, for θ ∈ (0, 1)

  Thus, the joint pf–pdf of Θ and *X* is

$$f_{\Theta X}(\theta, x) = f_{X|\Theta}(x|\theta) f_{\Theta}(\theta) = \frac{\theta^{\alpha+x-1}(1-\theta)^{(\beta-x+1)-1}}{B(\alpha,\beta)}.$$

  from which we compute the marginal pf of X by integration to obtain

$$f_X(x) = \int_0^1 \frac{\theta^{\alpha+x-1}(1-\theta)^{(\beta-x+1)-1}}{B(\alpha,\beta)} d\theta$$
$$= \frac{B(\alpha+x, \beta-x+1)}{B(\alpha,\beta)}$$

5 Bernoulli Random Variable

- Thus, we conclude

$$f_{\Theta|X}(\theta|x) = \frac{f_{\Theta X}(\theta, x)}{f_X(x)} = \frac{\theta^{\alpha+x-1}(1-\theta)^{(\beta-x+1)-1}}{B(\alpha+x, \beta-x+1)}$$

which is the pdf of a beta distribution with parameters $\alpha + x$ and $\beta - x + 1$.

5 Bernoulli Random Variable

- Example:
  In the previous example , if there is a sample of n observations of X denoted by X = $\{X_1, \ldots X_n\}$, compute the posterior pdf of Θ.

- Solution:
  We first compute the likelihood of X as follows

$$f_{X|\Theta}(x|\theta) = \prod_{i=1}^{n} \theta^{x_i}(1-\theta)^{1-x_i}$$

$$= \theta \sum_{i=1}^{n} x_i (1-\theta) \sum_{i=1}^{n} (1-x_i)$$

and the joint pf–pdf is

$$f_{\Theta X}(\theta, x) = f_{x|\Theta}(x|\theta) f_{\Theta}(\theta)$$

$$= \left[\theta^{\sum_{i=1}^{n} x_i} (1-\theta)^{\sum_{i=1}^{n}(1-x_i)}\right] \left[\frac{\theta^{\alpha-1}(1-\theta)^{\beta-1}}{B(\alpha,\beta)}\right]$$

$$= \frac{\theta^{(\alpha+n\bar{x})-1}(1-\theta)^{(\beta+n-n\bar{x})-1}}{B(\alpha,\beta)}$$

5 Bernoulli Random Variable

$$
\begin{aligned}
f_X(x) &= \int_0^1 f_{\Theta X}(\theta, x) d\theta \\
&= \int_0^1 \frac{\theta^{(\alpha+n\bar{x})-1}(1-\theta)^{(\beta+n-n\bar{x})-1}}{B(\alpha,\beta)} d\theta \\
&= \frac{B(\alpha+n\bar{x}, \beta+n-n\bar{x})}{B(\alpha,\beta)}
\end{aligned}
$$

we conclude that

$$
\begin{aligned}
f_{\Theta|X}(\theta|x) &= \frac{f_{\Theta X}(\theta, x)}{f_X(x)} \\
&= \frac{\theta^{(\alpha+n\bar{x})-1}(1-\theta)^{(\beta+n-n\bar{x})-1}}{B(\alpha+n\bar{x}, \beta+n-n\bar{x})}
\end{aligned}
$$

and the posterior pdf of follows a beta distribution with parameters $\alpha + n\bar{x}$ and $\beta + n - n\bar{x}$.

5 Bernoulli Random Variable

Note that the denominator in combined equation is a function of *x* but not θ.
Denoting

$$K(x) = \frac{1}{f_{\theta\epsilon\Omega_\Theta} f_{X|\Theta}(x|\theta) f_\Theta(\theta) d\theta}$$

we can rewrite the posterior pdf of Θ as

$$f_{\Theta|x}(\theta|x) = K(x) f_{x|\Theta}(x|\theta) f_\Theta(\theta)$$
$$\alpha f_{x|\Theta}(x|\theta) f_\Theta(\theta)$$

- K(x) is free of θ and is a constant of proportionality. It scales the posterior pdf so that it integrates to 1. The expression $f_{x|\Theta}(x|\theta) f_\Theta(\theta)$ enables us to identify the functional form of the posterior pdf in terms of θ without computing the marginal pdf of X.