第 15 卷 第 5 期
2004 年 9 月
水 科 学 进 展
ADVANCES IN WATER SCIENCE
Vol. 15, No. 5
Sep. , 2004

贝叶斯方法在洪水保险费调整中的应用

付 湘[1], 刘 宁[2], 王丽萍[1], 纪昌明[1]

(1. 武汉大学水资源与水电工程科学国家重点实验室, 湖北 武汉 430072; 2. 中华人民共和国水利部, 北京 100053)

摘要: 在洪水保险中, 获得的样本信息不符合对统计样本的理论要求。因此, 本文运用贝叶斯统计方法, 依据先验信息数据确定的保险费, 结合新的理赔记录, 调整和校正赔款频率和平均赔款额, 从而正确估计保险费。使其符合实际的风险水平。并运用实例分析了贝叶斯方法的可行性, 本文的方法和结论可供开展洪水保险项目研究和业务工作参考。

关 键 词: 洪灾损失; 贝叶斯方法; 洪水保险费

中图分类号: F840.64; TV122　　文献标识码: A　　文章编号: 1001-6791(2004)05-0675-04

1 贝叶斯方法

贝叶斯统计方法起源于英国学者贝叶斯(1702– 1761 年)的论文《论有关机遇问题的求解》。在此文中, 贝叶斯提出了著名的贝叶斯公式和一种归纳推理的方法。此后, 数学家拉普拉斯由贝叶斯方法导出了重要的相继律, 贝叶斯方法从 20 世纪 30 年代起逐渐发展为一个有影响的统计学派, 在工业、经济、管理等领域中得到广泛应用。

传统的数理统计方法对随机变量分布的估计是建立在具有独立性和代表性的样本信息基础上对随机变量分布参数的估计。但在洪水保险中, 往往难以获得足够的样本信息, 或仅有的理赔记录不符合对统计样本的理论要求。这时, 对随机变量分布的估计就需要掺入评估人的主观判断, 并利用新获得的证据来修正原来的估计。

设随机变量 X 的分布类型为 $F(x, \theta)$, 在连续情形下相应的密度函数族为 $f(x, \theta)$, 估计 θ 参数的贝叶斯法与经典的数理统计法的基本区别就是把参数 θ 看作为随机变量且服从某一概率分布。因而可记作 θ', 这样 θ 本身应有一个概率分布。在此假定下, 参数估计的贝叶斯方法可概括为以下步骤:

步骤 1 选择先验分布 设 θ 的分布函数和密度函数分别为 $F(\theta)$ 和 $f(\theta)$, 称为先验分布和先验密度, 它反映了评估者对参数 θ 的情况有一个初步的看法或信念。

步骤 2 确定似然函数 评估人针对随机变量 X 进行了一些试验或观察以获得一些新的信息, 假设所获得的观察值为 $x_1, x_2, \cdots, x_n$, 则在 $\theta = \theta$ 的假定下, 可构造似然函数

$$f(x \mid \theta) = L(x_1, x_2, \cdots, x_n; \theta) = \prod_{i=1}^{n} f(x_i \mid \theta) \qquad i = 1, 2, \cdots, n \tag{1}$$

步骤 3 确定参数 θ 的后验分布 按照关于条件概率的贝叶斯定理, 可以求得关于参数 θ 的后验分布, 对密度而言, 有

$$f(\theta \mid x) = \frac{f(x \mid \theta) f(\theta)}{\int f(x \mid \theta) f(\theta) \mathrm{d}\theta} \tag{2}$$

步骤 4 选择损失函数 引入一个非负函数, 记作 $Loss(\hat{\theta}, \theta)$ 来刻画参数的真实值 θ 与估计值 $\hat{\theta}$ 之间差距

收稿日期: 2003-06-06; 修订日期: 2003-09-02
基金项目: 国家自然科学青年基金资助项目 (50209011); 武汉大学科技创新基金; 河海大学水资源开发教育部重点实验室开放基金资助项目
作者简介: 付 湘(1971–), 女, 江西九江人, 武汉大学副教授, 博士, 主要从事水文及水资源教学与研究工作。
E-mail: xfu@wuhee.edu.cn

的严重程度，这个函数俗称损失函数。

步骤5 估计参数 根据所选择的损失函数和参数的后验分布，通过求损失函数的期望值的最小值的解来作为参数 θ 的贝叶斯估计值[1]。即求解：

$$\min_{\hat{\theta}} E[Loss(\hat{\theta}, \theta)] \min_{\hat{\theta}} \int_{-\infty}^{+\infty} Loos(\hat{\theta}, \theta) f(\theta | x) \mathrm{d}\theta \tag{3}$$

这里选择平方损失函数 $Loss(\hat{\theta}, \theta) = (\hat{\theta} - \theta)^2$，代入上式求解得最优解 $\hat{\theta} = E(\theta | x)$。 (4)

与数理传统统计法相比，贝叶斯法明确地认可研究者的主观判断。参数估计的贝叶斯法可分为两个阶段，第一阶段是根据所选择的先验概率并利用新的观察信息求出后验概率；第二阶段是根据所选择的损失函数并利用求最小平均损失来求未知参数的贝叶斯估计。第一阶段的主观性体现在对先验分布的选择上，第二阶段的主观性则体现在对损失函数的选择上。

本文贝叶斯法主要用于保险费的调整。即根据先验信息数据确定的保险费，结合新的理赔记录，调整和校正赔款频率和平均赔款额，从而正确估计风险保费。

2 洪水保险费的初定

洪水保险费的初步确定可用下式：

保险费= 洪水保险损失(或纯保费)+ 安全费+ 附加费。

其中，洪水保险损失是根据一定时期保险赔款总额确定，可作为制定纯保费的基本指标，常用多年平均洪灾损失表示。安全费是为了消除保险计算中的不确定性而增加的。实际计算中可按纯保费的一定比值来确定安全费。附加费是以保险人经营保险业的各种营业费用和保险利润为基础的，其费用于保险人的营业费用支出和提供部分保险利润。洪水保险的利润率通常很低，或实行非盈利经营。附加费的计算也可按纯保费的一定比例确定。

根据以上的分析，洪水保险费的计算公式为：

$$P_f = NP_f + (\lambda_1 + \lambda_2) NP_f = (1 + \lambda_1 + \lambda_2) NP_f \tag{5}$$

式中 P_f 为洪水保险费，也称毛保险费；NP_f 为纯保险费；λ_1 为安全系数，λ_2 为附加系数。从式(5)可看出，厘定洪水保险费的核心是确定纯保费，亦即多年平均损失。只要纯保费计算出来了，再加上适当的比例就可确定出洪水保险的毛保费。

洪灾损失按时间尺度可分为各次灾害损失、年度灾害损失和多年平均灾害损失等，为确定洪水保险费所依据的灾害损失，需要具有较好的稳定性，因此一般采用多年平均损失。多年平均洪灾损失的计算分为两部分：首先，运用典型地区抽样调查方法调查当地以往的灾害损失情况，运用线性回归分析法或等级相关关系法，建立各类资产的损失率与洪水要素的相关关系[2]。然后，应用分类财产洪灾损失率关系，对各种频率的洪水，计算评价区域各类财产的总损失。

此外，在应用洪水实测资料进行分析计算时，需先对资料进行可靠性、一致性与代表性分析。排除资料中可能存在的错误，审查水文现象影响因素是否一致以及资料对于水文变量总体的代表性。

3 洪水保险费的调整

洪水保险的风险保费(即通常所说的纯保险费)一般是通过估计索赔频率和平均索赔额来计算[3]。所谓索赔频率是指每个风险单位在保险责任期内的索赔次数，而平均索赔额就是平均每个风险单位的损失额。根据等价原理，风险保费应该为这两者的乘积。现考虑 n 份保险单组合的索赔记录，若每份保险单的预期赔款频率为 $\bar{q}$，在某年中赔款次数是均值为 $\bar{q}$ 的随机变量，如果每份保险单每次赔款额是均值为 $\bar{m}$、方差为 σ^2 的独立随机变量，那么全年每份保险单的风险保费的均值为 $\bar{q} \cdot \bar{m}$，保险人需要的是对风险保费 $P = \bar{q} \cdot \bar{m}$ 的估计。

3.1　索赔频率的校正

如果根据先验信息，索赔频率 q 服从 $\Gamma(\alpha,\ \beta)$，而在给定 $q=0$ 的条件下，每份保单的索赔次数 X 服从泊松分布 $P(\theta)$。于是，根据贝叶斯公式，q 的后验密度函数为

$$f(\theta \mid x_1,\ x_2,\ \cdots,\ x_n)=\frac{\beta+n}{\Gamma\left(\alpha+\sum_{i=1}^{n}x_i\right)}\exp\left(-(\beta+n)\theta[(\beta+n)\theta]^{\alpha+\sum_{i=1}^{n}X_i-1}\right) \tag{6}$$

这是以 $\alpha+\sum_{i=1}^{n}x_i$、$\beta+n$ 为参数的 Γ 分布密度函数。所以，q 的贝叶斯估计为

$$\hat{q}_B=E(q \mid x_1,\ x_2,\ \cdots,\ x_n)=\frac{\alpha+\sum_{i=1}^{n}x_i}{\beta+n}=\frac{\beta}{\beta+n}\cdot\frac{\alpha}{\beta}+\frac{n}{\beta+n}\cdot\frac{\sum_{i=1}^{n}x_i}{n} \tag{7}$$

式中 $\sum_{i=1}^{n}x_i/n$ 是通过观察得到的 n 份保单的平均索赔次数，$\hat{q}_B$ 就是利用后者对前者进行校正。

3.2　平均索赔额的校正

平均索赔额也称为平均损失额，记为 m。如果索赔额在 $m=\theta$ 的条件下，服从 $N(\theta,\ \sigma_1^2)$，而参数 m 也是个服从 $N(\mu,\ \sigma_2^2)$，根据贝叶斯公式 m 的后验密度函数为

$$f(\theta \mid x_1,\ x_2,\ \cdots,\ x_n)=\frac{1}{\sqrt{2\pi}\sqrt{\frac{\sigma_1^2\sigma_2^2}{\sigma_1^2+n\sigma_2^2}}}\exp\left\{-\left(\theta\frac{\mu\sigma_1^2+\sigma_2^2\sum_{i=1}^{n}x_i}{\sigma_1^2+n\sigma_2^2}\right)^2\Big\backslash\frac{2\sigma_1^2\sigma_2^2}{\sigma_1^2+n\sigma_2^2}\right\} \tag{8}$$

这是服从 $N\left(\left(\mu\sigma_1^2+\sigma_2^2\sum_{i=1}^{n}x_i\right)\Big\backslash(\sigma_1^2+n\sigma_2^2),\ \frac{\sigma_1^2\sigma_2^2}{\sigma_1^2+n\sigma_2^2}\right)$ 的密度函数，所以 m 的贝叶斯估计为

$$\hat{m}_B=E(m \mid x_1,\ x_2,\ \cdots,\ x_n)=\frac{\mu\sigma_1^2+\sigma_2^2\sum_{i=1}^{n}x_i}{\sigma_1^2+n\sigma_2^2}=\frac{\mu\sigma_1^2}{\sigma_1^2+n\sigma_2^2}+\frac{n\sigma_2^2}{\sigma_1^2+n\sigma_2^2}\frac{\sum_{i=1}^{n}x_i}{n} \tag{9}$$

式中 $\sum_{i=1}^{n}x_i/n$ 是通过观察得到的平均索赔额；$\hat{m}_B$ 就是利用后者对前者进行校正。

4　实例分析

某保险公司新开洪水保险业务，以该地区洪水灾害发生的平均频率 0.148 作为先验信息(即 $\alpha/\beta=0.148$)，估计多年平均损失额 $\mu=4\,000$ 元，$\sigma_1=1\,000$ 元，$\sigma_2=20\,000$ 元。该保险公司的负责人根据经验，有相当把握(95%)认为真实赔款频率与 0.148 的相对误差不会超过 25%。结果，第一个业务年度该险种的 2 427 份保单，发生了 320 件赔案，平均索赔额为 4 500 元。在第二个经营业务年度中，有效保单为 6 982 份，发生赔案 951 件，平均索赔额为 4550 元。保险公司利用上述贝叶斯方法对索赔频率和平均索赔额进行校正。

(1) 索赔频率的校正

根据保险公司负责人的经验
$$P\left(\left|\left(q-\frac{\alpha}{\beta}\right)\Big\backslash\sqrt{\frac{\alpha}{\beta^2}}\leqslant 25\%\sqrt{\alpha}\right|\right)=95\% \tag{10}$$

由中心极限定理可知，$U=q-\frac{\alpha}{\beta}\Big\backslash\sqrt{\frac{\alpha}{\beta^2}}$ 近似地服从标准正态分布，所以 $0.25\sqrt{\alpha}=U_{p=97.5\%}=1.96\approx 2$，由此得：$\alpha=64$，则 $\beta=\alpha/0.148=432.5$。

因此，根据第 1 年的经营情况，索赔频率可校正为：$\hat{q}_B=0.134$；又根据第 2 年的经营情况，索赔频率可进一步校正为：$\hat{q}_B=0.136$。

则根据两年的经营情况校正后的索赔率比先验索赔频率0.148 降低了0.012。

(2) 平均索赔额的校正

根据第1 年的经营情况, 平均赔款额可校正为: $\hat{m}_B = 4222$ 元; 又根据第2 年的经营情况, 平均赔款额可进一步校正为: $\hat{m}_B = 4410.2$ 元。

该公司根据索赔频率和平均索赔额的先验信息厘定的风险保费为 $0.148\times4000=592$ 元, 而经过两个业务年度的实践, 利用校正后的索赔频率和平均索赔额, 风险保费应调整为 $0.136\times4410.2=600$ 元。

随着数据的不断积累, 真实赔款频率和平均索赔额的确定性将越来越好, 由此厘定的保费将更趋于合理。

5 小 结

本文中贝叶斯方法主要用于保险费的调整。即根据先验信息数据确定的保险费, 结合新的理赔记录, 调整和校正赔款频率和平均赔款额, 从而正确估计风险保费。

(1) 传统的数理统计法对随机变量分布的估计是建立在具有独立性和代表性的样本信息基础上对随机变量分布参数的估计。但在洪水保险中, 往往难以获得足够的样本信息, 或仅有的理赔记录不符合统计样本的理论要求。这时, 对随机变量分布的估计就需要运用贝叶斯法, 即掺入评估人的主观判断, 并利用新获得的证据来修正原来的估计。其缺点是必须已知先验分布, 且损失函数需选择平方损失函数以得到信度因子。

(2) 厘定洪水保险费的核心是确定纯保费, 亦即多年平均损失。它的初步计算分为两部分: 首先, 运用典型地区抽样调查方法调查当地以往的灾害损失情况, 并对资料进行可靠性、一致性与代表性分析; 运用线性回归分析法或等级相关关系法, 建立各类资产的损失率与洪水要素的相关关系。然后, 应用分类财产洪灾损失率关系, 对各种频率的洪水, 计算评价区域各类财产的总损失。

(3) 由实例分析可知, 保险公司根据索赔频率和平均索赔额的先验信息厘定的风险保费为592 元, 而经过两个业务年度的实践, 风险保费调整为600 元。在市场经济中, 增加风险保费可降低赔付率, 但是会导致投保人望而生畏。因此, 增加风险保费的幅度不能过大, 并辅以增加免赔率以降低赔付率。

参考文献:

[1] James O, Berger. 统计决策论及贝叶斯分析[M]. 贾乃光译. 北京: 中国统计出版社, 1998. 139-180.

[2] 刘树坤, 宋玉山, 程晓陶, 等. 黄河滩区及分滞洪区风险分析和减灾对策[M]. 郑州: 黄河水利出版社, 1999. 141-179.

[3] 谢志刚, 韩天雄. 风险理论与非寿险精算[M]. 天津: 南开大学出版社, 2000. 277-283.

Application of Bayesian approach to the adjustment of flood premium*

FU Xiang[1], LIU Ning[2], WANG Li-ping[1], JI Chang-ming[1]

(1. *State Key Laboratory of Water Resources and Hydropower Engineering Science*, *Wuhan University*, *Wuhan* 430072, *China*;

2. *Ministry of Water Resources of the People's Republic of China*, *Beijing* 100053, *China*)

Abstract: In the flood insurance, the sample information gotten doesn't accord with the theory requset for the statistical sample. So the Bayesian approach is used, namely according to premium confirmed by priori information, claim frequency and expected claim size are adjusted and corrected combining the new claims settlement record. Thus premium can be estimated correctly and suits actual risk level. And an example is used to check the feasibility of Bayesian approach. The method and conclusion of the paper can afford reference for project research and business of flood insurance.

Key words: expected flood damage; Bayesian approach; flood premium

* The project is supported by the National Science Young Natural Foundation of China (No. 50209011).