信度估计的理论与方法

温利民　著

科学出版社

信度估计的理论与方法

温利民　著

科　学　出　版　社
北　京

内 容 简 介

信度理论是基于样本索赔数据和先验信息对保费定价的一种方法。本书以贝叶斯统计为基础，利用各种统计方法，对信度估计的理论与方法作了系统和详细的介绍。从保费原理与信度、相依风险与信度等方面探讨了现代信度理论方法的前沿研究课题。

本书可以作为保险精算、金融数学、概率论数理统计专业的教学或参考用书，同时也适合金融、保险、统计等行业的从业人员，特别是正从事保险精算和金融创新设计的决策人员以及相关的研究工作者阅读参考。

图书在版编目（CIP）数据

信度估计的理论与方法/温利民著. —北京：科学出版社，2012
ISBN 978-7-03-035597-3

Ⅰ. ①信… Ⅱ. ①温… Ⅲ. ①财产保险-保险费-信度-研究

Ⅳ. ①F840.65

中国版本图书馆 CIP 数据核字 (2012) 第 222942 号

责任编辑：唐 薇 马 跃／责任校对：林青梅
责任印制：阎 磊／封面设计：迷底书装

科学出版社 出版
北京东黄城根北街 16 号
邮政编码：100717
http://www.sciencep.com

中国科学院印刷厂印刷

科学出版社发行 各地新华书店经销

*

2012 年 10 月第 一 版 开本：B5(720 × 1000)
2012 年 10 月第一次印刷 印张：13
字数：267 000

定价：52.00 元

(如有印装质量问题，我社负责调换)

前　　言

保费定价，也称保费厘定，是精算师根据保单的风险状况为保单制定合适的保费的过程。在厘定保费时，保险公司最关心的问题主要有两个：一是如何使征收的保费足够理赔；二是在保费足够理赔的基础上，如何增强保险产品的竞争力。合适的保费厘定是保险公司稳健经营的前提，也是保险产品具有竞争力的关键因素。在对保险产品定价的过程中，精算师必须对风险进行科学的分析和评价，以使制定的价格能确保保险公司承担正常的赔付和获取一定的利润。征收的保费过高，会使得投保人数量减少，保单流失；过低，会使得征收的保费不够理赔，造成亏损。保险公司在厘定费率时，既要考虑到总的保费收入，又要考虑到投保人的预期保费，使得保费在投保人之间公平分摊。

在精算领域，保费厘定最重要的方法之一就是信度原理。信度原理是基于贝叶斯 (Bayes) 框架，结合被保险人的索赔样本数据信息和先验分布信息对保单定价的一种方法。本书以贝叶斯统计为基础，利用数理统计中的方法，结合金融研究工具介绍信度估计的理论与方法，构建各种风险模型下的保费定价模型。

由于信度理论的方法是建立在贝叶斯框架下的，因此本书的阅读者应对贝叶斯统计有所了解。

本书分为四个部分：第一部分是信度理论综述，主要介绍保费、保费原理，以及信度理论的主要文献，对目前信度理论的研究方向作了概括。第二部分包括三章，主要介绍了经典的信度理论研究方法。通过引进贝叶斯理论的基本研究框架，详细介绍了经典的 Bühlmann 信度理论、多元信度理论和回归信度理论。第三部分是对信度理论与保费原理的讨论与研究。由于传统的信度理论都只基于净保费原理建立模型，而净保费不具有安全负荷性，所以在实际运用中，保险公司要选取具有安全负荷的合适的保费原理。本部分在 Esscher 保费原理、指数保费原理、广义加权保费原理下建立了信度估计模型，研究了信度估计的性质。最后，基于损失的分布函数的信度模型，建立了统一的保费厘定，统一了大部分保费原理下的信度模型。第四部分是对相依风险与信度理论的探讨，在共同效应相依、一般风险相依模型下建立了信度估计理论，研究了估计的性质。

本书内容是基于信度理论的最新文献进行的拓广研究，力图把理论和实际运用相结合，涉及数学方面的内容都有详细的证明，尽量做到通俗易懂。

本书的出版得到国家自然科学基金 (71001046，71063006) 和江西省自然科学基金 (20114BAB211004) 的资助，在此表示感谢。特别感谢我的博士导师王静龙老

师、吴贤毅老师，是他们的关心、鼓励和帮助，本书才得以顺利完成。同时，我还要感谢科学出版社的编辑同志，是他们的大力支持和辛勤劳动，才使得本书得以顺利出版。

温利民
江西师范大学
2012 年 9 月

目　　录

第四篇 风险相依与信度理论

第一篇

信度理论综述

第 1 章　保费定价与保费原理

保费定价是指精算师对保险产品制定一个合理的价格的过程。在厘定保费时，保险公司最关心的问题是：① 如何使征收的保费足够理赔；② 在保费足够理赔的基础上，如何增强保险产品的竞争力。在对保险产品定价过程中，精算师必须对风险进行科学的分析和评价，以使制定的价格能确保保险公司正常的赔付和获取一定的利润。征收的保费过高，会使投保人数量减少，保单流失；过低，会使征收的保费不够理赔，造成亏损。因此，保费定价是保险公司最为关注的重要问题之一。保险公司在厘定费率时，既要考虑总的保费收入，又要考虑投保人的预期保费，使得保费在投保人之间公平分摊。

通常情况下，精算师制定保费的依据是保险产品的历史索赔数据。在精算学中，把一份保单可能导致的索赔定义为一个风险，用随机变量 X 来表示。这时该保险的历史索赔数据可以看做该随机变量的随机样本的实现值。通过分析和了解这些数据信息，得到风险随机变量的分布特征，进而为该风险 (保单) 制定合理的价格 $H(X)$，即为保费。

定义 1.0.1　*设 X 是取值非负的风险随机变量，其分布函数为 $F_X(x)$，保费原理就是给风险 X 分配一个实值泛函 $H(\cdot)$，记为 $X \to H(X)$ 或 $F_X \to H(F_X)$。*

通俗地讲，保费原理就是为取值随机的风险变量，制定一个非随机的价格。事实上，保费原理就是一种风险度量工具，度量了风险 X 的大小。在这个意义上，保费原理一般要求满足风险度量的一致性公理 (Artzner et al., 1999)。

在实际运用中，常用的保费原理如下。

1. 净保费原理

$$H(X) = E(X)$$

2. 期望值原理

$$H(X) = (1+\alpha)E(X)$$

式中，$\alpha > 0$ 为常数。

3. 方差原理

$$H(X) = E(X) + \alpha \mathrm{Var}(X)$$

式中，$\alpha > 0$ 为常数。

4. 修改方差原理

$$H(X) = E(X) + \alpha \frac{\mathrm{Var}(X)}{E(X)}$$

式中，$\alpha > 0$ 为常数。

5. 标准差原理

$$H(X) = E(X) + \alpha\sqrt{\mathrm{Var}(X)}$$

式中，$\alpha > 0$ 为常数。

6. 指数原理

$$H(X) = \frac{1}{\alpha}\log E[\exp(\alpha X)]$$

式中，$\alpha > 0$ 为常数。

7. Esscher 原理

$$H(X) = \frac{E\left(X\mathrm{e}^{\alpha X}\right)}{E\left(\mathrm{e}^{\alpha X}\right)}$$

式中，$\alpha > 0$ 为常数。

8. Kamp 保费

$$H(X) = \frac{E\left[X\left(1-\mathrm{e}^{-\alpha X}\right)\right]}{E\left(1-\mathrm{e}^{-\alpha X}\right)}$$

式中，$\alpha > 0$ 为常数。

9. 条件尾期望原理

$$H(X) = E(X|X>q) = \frac{E[XI(x>q)]}{P(x>q)}$$

式中，$q > 0$ 为常数。

10. 修正条件尾期望原理

$$H(X) = E(X|X>q) + \alpha\frac{\mathrm{Var}(X|X>q)}{E(X|X>q)}$$

式中，$q > 0, \alpha > 0$ 为常数。

11. 失真保费原理

$$H(X) = \int_0^{\infty} g\left[S_X(x)\right]\mathrm{d}x$$

式中，函数 $g(\cdot)$ 为非降失真函数，满足 $g(0) = 0, g(1) = 1$。

12. 风险调整保费原理

$$H(X) = \int_0^{\infty}\left[S_X(x)\right]^{\frac{1}{p}}\mathrm{d}x$$

式中，$p > 1$ 为常数。

13. 零效用保费原理

保费 H 为方程

$$U(0) = E\left[U(H-X)\right]$$

的解，式中，U 为效用函数。

14. 分位数保费原理

$$H(X)=\min\{p, F_X(p)\geqslant 1-\varepsilon\}$$

式中，ε 为给定的小概率。

15. 绝对偏差保费原理

$$H(X)=E(X)+\alpha\kappa_X$$

式中，$\kappa_X=E\left|X-F_X^{-1}\left(\dfrac{1}{2}\right)\right|$ 为 X 的绝对中位数离差；$\alpha>0$ 为常数。

在理论上，若能根据样本很好地估计风险随机变量 X 的分布 F_X 或分布的某些特征 (如一二阶矩)，则保险公司可以选取合适的保费原理，得到该风险的价格估计，称为保费估计。

例 1.0.1　设 $X_1, X_2, \cdots, X_n$ 是风险保单 X 的 n 次观测值，假设为独立同分布的样本，具有共同的分布函数 $F_X(x)$。记 $F_n(x)$ 为样本 $X_1, X_2, \cdots, X_n$ 的经验分布函数，则根据 Borel-Cantelli 引理，有

$$\sup_{x\geqslant 0}\left|F_n(x)-F_X(x)\right|\longrightarrow 0,\ \text{a.s.}$$

若保险公司选取的保费是 Esscher 保费原理，则 $\widehat{H(X)}=\dfrac{\sum_{i=1}^{n}X_i\exp(\alpha X_i)}{\sum_{i=1}^{n}\exp(\alpha X_i)}$是保费 $H(X)=\dfrac{E[X\exp(\alpha X)]}{E[\exp(\alpha X)]}$ 的一个相合估计，即当样本容量 $n\longrightarrow\infty$ 时，有 $\widehat{H(X)}\to H(X)$, a.s.。

然而，在实际中，对某种个体保单的数据量太少而不足以使保费估计具有良好的性质。这时一个可取的办法就是利用相同类型的保单索赔数据，因为保险公司可能已经积累了大量相同类型保单的索赔数据，若把这些数据集中起来，则能大大增加样本容量。但这种保单索赔数据的融合必须基于两条假设：① 其他保单的索赔样本与原样本相互独立；② 其他保单的索赔样本与原样本具有相同的母体分布。若这两个假设成立，则称这些保单组合是齐次的，可以把这些保单组合的数据融合在一起对该保单定价。然而，在实际中，不同保单导致的索赔之间一般不具有共同的分布，甚至可能不是相互独立的，因为影响保单索赔的因素是非常复杂的。以汽车保险为例，车辆类型、司机的年龄和性别、汽车的行驶区域等都影响了保单的潜在索赔。因此，不同保单之间的风险常常是非齐次的。这就引发了一个问题：如何为非齐次的保险制定合理的价格？这个问题正是信度理论研究的核心课题。

第 2 章　信度理论介绍

2.1　信度理论解决的精算问题

信度理论是精算学中最重要的经验保费厘定技巧。它是一种经验估费模型。在这个过程中，精算师根据过去的单个风险或者一个保单组合风险的经验数据，调整未来的保险费。一般地，我们的目标是制定某个个体风险 (保险合同) 的保费，但是这个保险合同的索赔数据较少，而类似风险的保单组合却有较多的数据，这时是否能结合类似保险合同的数据从而制定该个体保险合同的保费？下面以一个例子说明信度理论提出的背景及意义。

假设某保单组合有 K 个独立风险 $X_i, i=1,2,\cdots,K$，由这 K 个风险导致的总索赔为 $S=\sum\limits_{i=1}^{K} X_i$。记 $H(\cdot)$ 为保险公司的保费原理，则该保单组合的总保费为 $H(S)$，现在的问题是：如何将总保费分摊到各个保单， 即每份保单应缴纳多少保费？若该保单组合风险是齐次的，则显然每份保单缴纳相同的保费 $\dfrac{1}{K}H(S)$是合理的，但看下面的例子。

例 2.1.1　设保单组合中有 12 份相同类型的保单，每份保单每年缴付相同的保费 $m=0.25$。假设每份保单产生的索赔为 Bernoulli 变量，即取值为 0 或 1。经过 10 年后，得到索赔数据的记录如表 2.1 所示。

表 2.1　保单组合 10 年的索赔记录

索赔 X_{ij} 保单号 i 年数 j	1	2	3	4	5	6	7	8	9	10	11	12
1		1	1						1		1	
2						1				1		
3			1				1		1			
4						1	1		1			
5											1	
6						1			1			
7					1				1			
8						1	1		1			
9						1					1	
10					1				1			
$\overline{x_i}$	0	0.2	0.2	0	0.2	0.5	0.3	0	0.7	0.1	0.3	0

表 2.1 中最后一行 $\overline{x_i}$ 表示各保单在 10 年内的索赔均值。根据表中的数据，容易算出平均每份保单每年的索赔成本为 $\dfrac{25}{120} \approx 0.21 < 0.25$。因此，保险公司可以保证偿付并获取一定的利润。但是，若去掉保单 6、保单 9，则平均每份保单每年的索赔成本为 $\dfrac{13}{120} \approx 0.11$，风险将大大改善，利润将显著提高。 若保险公司不采取任何措施，继续收取相同的保险费 $m = 0.25$，则司机 1、4、8、12 因为具有良好的驾驶记录而寻找能为他们提供较低费率的保险公司。若是这样的话，平均索赔成本将变为 $\dfrac{25}{80} \approx 0.3125 > 0.25$，则保险公司将亏本经营。当然，除了对每个保险合同收取相同保费 m 之外，还有另一种厘定保费的方法，即对不同的保险合同根据他的索赔情况收取不同的保费。当然，保险公司可能认为应该相信数据，决定对各保单今后分别收取的保费为这些年各保单的索赔个体均值 $\overline{x_i}$。这时司机 1、4、8、12 将因为不需要缴纳任何保费而选择留在该保险公司，而司机 6、9，甚至司机 7、11 等因缴纳的保费太高而离开，保险公司将失去大量保单。事实上，由于个体索赔数据较少，索赔较多的现象可能是由驾驶的偶然因素导致。由这两种不同的厘定保费的方法所引起的保单流失和亏本经营等情况都不是保险公司希望看到的。分析其原因，是保单风险之间的非齐次性导致的，这可由经典的统计理论进行证明。

设保单 i 在第 j 年的索赔为 $X_{ij} \sim B(1, \theta_i), j = 1, 2, \cdots, 10$。风险齐性假设为

$$H_0 : \theta_1 = \theta_2 = \cdots = \theta_{12}$$

检验统计量为

$$\chi^2 = \frac{10\displaystyle\sum_{i=1}^{12}\left(\widehat{\theta}_i - \widehat{\theta}\right)^2}{\widehat{\theta}\left(1 - \widehat{\theta}\right)} \overset{H_0}{\sim} \chi^2(11) \tag{2.1}$$

式中，$\widehat{\theta}_i = \overline{X_i}$；$\widehat{\theta} = \overline{\overline{X}} = \dfrac{1}{12}\displaystyle\sum_{i=1}^{12}\overline{X_i}$。取检验水平 $\alpha = 0.05$，查表得 $\chi^2_{0.95}(11) = 19.675$。因为此时统计量的值 $\chi^2 = 32.1 > \chi^2_{0.95}(11)$，则根据假设检验的原理，拒绝原假设，即认为该保单组合的风险是非齐次的。

从上面的分析结果看，保险公司无论是根据保单合同历史数据的均值定价，还是继续实行以前的聚合保费定价方法，都将出现不良后果。因此，在对这种具有非齐次性的保单定价时，传统的保费定价方法面临着大的挑战。解决这种非齐次风险保单定价问题，主要是采用信度原理。信度理论不仅利用个体保单的索赔数据，而且利用相同类型的保单组合数据，定价的保费为两者的加权平均：

$$\text{信度保费} = Z \times \text{个体均值} + (1 - Z) \times \text{聚合均值} \tag{2.2}$$

式中，$0 \leqslant Z \leqslant 1$ 被称为信度因子，是样本容量 n 的某个函数。当个体数据的样本容量增大时，Z 趋于 1，它反映了个体样本数据的可信度。

信度理论自 20 世纪初由精算师作为一种直观想法提出后，许多学者在贝叶斯框架下研究该问题，并得到了统计意义上较好的解释。在理论与实践相结合的过程中，发展并形成了今天的经验厘定信度原理。至今为止，信度理论已经拥有庞大的分支体系，为保险业提供了强大的理论支撑。

2.2 关于信度原理的历史评注

从 20 世纪初到现在，信度理论的研究主要经历了两个发展阶段，更确切地说是两个不同的分支：① 建立在频率方法上的有限扰动理论 (limited fluctuation)；② 以贝叶斯理论为基础的最大精确度信度理论 (greatest accuracy credibility)。这两种方法都是希望通过已有的历史数据来合理地制定保费。

有限扰动理论是 Mowbray 于 1914 年在对工伤补偿保险 (workers' compensation) 进行保费厘定时提出的，而 Stellwangen (1925) 将有限扰动理论应用于汽车保险，Perryman (1932) 对有限扰动理论进行了解释。最大精确度信度理论也称为欧式信度理论。最早提出最大精确度信度理论的是 Whitney (1918)。他在研究工伤补偿保险的保费厘定时，假设一个雇主拥有 P 个雇员，总索赔服从二项分布 $B(P, M)$，而索赔概率 M 本身也是正态分布随机变量，事实上这正是贝叶斯框架下的二项–正态模型。根据贝叶斯定理可以估计参数 M，得到下一期的保费估计具有信度加权的形式。随后 Keffer (1929) 将最大精确度信度原理运用到群体寿险 (group life insurance)。Bailey (1945) 运用最小二乘法对最大精确信度理论建立数学模型，得到正态–正态、Beta–二项等分布模型下的信度估计。真正无分布 (任意分布) 的最大精确度信度理论则是 Bühlmann (1967) 建立的。他仍然用最小二乘法，在贝叶斯框架下，将估计限定在样本的线性函数类中，得到的最优保费估计恰好为信度加权形式，从而建立了无分布信度理论。

2.2.1 有限扰动理论

有限扰动信度理论是用于确定完全信度需要的最小样本容量的方法。设 $\widehat{\theta}_i$ 为 θ_i 的完全信度估计，即存在某个临界样本容量 n_0，使得 $\widehat{\theta}_i$ 落入 θ_i 的 $100k\%(k>0)$ 置信区间内的概率至少为 $1-\varepsilon$。该问题可简化为求最小样本容量 n，使得

$$\Pr\left(\left|\widehat{\theta}_i - \theta_i\right| \leqslant k\theta_i\right) \geqslant 1-\varepsilon \tag{2.3}$$

利用正态近似很容易求出最小样本容量 n_0。

例 2.2.1　设 $\varPhi(\cdot)$ 表示标准正态分布的分布函数。设 $X_{ij} \overset{i.i.d}{\sim} b(1,\theta_i), j = 1,2,\cdots,n$。用 $\overline{X}_i$ 估计 θ_i。给定的 $k>0, \varepsilon>0$，因为

$$\Pr\left(\left|\widehat{\theta}_i-\theta_i\right| \leqslant k\theta_i\right)=\Pr\left[\frac{\sqrt{n}\left|\widehat{\theta}_i-\theta_i\right|}{\sqrt{\theta_i\left(1-\theta_i\right)}} \leqslant \frac{\sqrt{n}k\theta_i}{\sqrt{\theta_i\left(1-\theta_i\right)}}\right]=2\varPhi\left[\frac{\sqrt{n}k\theta_i}{\sqrt{\theta_i\left(1-\theta_i\right)}}\right]-1$$

则满足式 (2.3) 的最小样本容量为

$$n_0=\frac{U_{1-\frac{\varepsilon}{2}}^2\left(1-\theta_i\right)}{k^2\theta_i}$$

若取 $\varepsilon=0.1, k=0.05, \theta_i=0.5$，则 $n_0 \approx 1082$，即当样本容量 $n>1082$ 时，可以完全相信样本，即 $\widehat{\theta}_i=\overline{X_i}$。

若 $n<n_0$，则完全信度条件不满足，这时就必须考虑部分信度。对 $0 \leqslant Z \leqslant 1$，令

$$\widehat{\theta}_i^{\,a}=Z\overline{X}_i+(1-Z)\,\overline{\overline{X}} \tag{2.4}$$

式中，$\overline{\overline{X}}$ 为聚合均值。这时估计误差为

$$\widehat{\theta}_i^{\,a}-\theta_i=Z\left(\overline{X}_i-\theta_i\right)+(1-Z)\left(\overline{\overline{X}}-\theta_i\right)$$

式中，$\overline{X}_i-\theta_i$ 描述了用个体均值 $\overline{X}_i$ 估计 θ_i 时的误差。有限扰动理论要求该误差的绝对值小于 $k\theta_i$ 的概率大于或等于 $1-\varepsilon$，即

$$\Pr\left(Z\left|\overline{X_i}-\theta_i\right| \leqslant k\theta_i\right) \geqslant 1-\varepsilon$$

令

$$\begin{aligned}
&\Pr\left(Z\left|\widehat{\theta}_i-\theta_i\right| \leqslant k\theta_i\right)\\
=&\Pr\left[\frac{\sqrt{n}Z\left|\widehat{\theta}_i-\theta_i\right|}{\sqrt{\theta_i\left(1-\theta_i\right)}} \leqslant \frac{\sqrt{n}Zk\theta_i}{\sqrt{\theta_i\left(1-\theta_i\right)}}\right]\\
=&2\varPhi\left[\frac{\sqrt{n}Zk\theta_i}{\sqrt{\theta_i\left(1-\theta_i\right)}}\right]-1\\
=&1-\varepsilon
\end{aligned}$$

则

$$Z^2=\frac{k^2\theta_i n}{U_{1-\frac{\varepsilon}{2}}^2\left(1-\theta_i\right)}=\frac{n}{n_0}$$

即 $Z=\sqrt{\dfrac{n}{n_0}}$。由于 $0 \leqslant Z \leqslant 1$，则可取 $Z=\min\left(\sqrt{\dfrac{n}{n_0}},1\right)$。当 $Z<1$ 时称为部分信度，当 $Z=1$ 时称为完全信度。

2.2.2 最大精确度信度理论

从 Whitney (1918) 提出二项 – 正态分布模型下的最大精确度信度原理到 Bühlmann (1967) 用最小二乘法建立无分布信度理论，整整经历了半个世纪。从此，最大精确度信度理论的发展进入了快车道，该理论引起的研究兴趣及其在保险中的成功运用早已超过了有限扰动理论。20 世纪 70~80 年代可看做是信度理论的全面发展时期，20 世纪 90 年代以来，信度理论逐步完善。至今，信度理论的定价研究仍然是精算界研究的热点问题。这里，对 Bühlmann 的经典信度理论基本想法与结论作简单介绍，具体可参考 Bühlmann (1967；1969)、Bühlmann 和 Straub (1970) 的研究或 Bühlmann 和 Gisler (2005) 的信度理论专著。

设 $X_1, X_2, \cdots, X_n$ 为某个风险 (保单) 前 n 年的历史索赔记录，该保单风险的潜在损失 X 由风险参数 θ 识别，当风险参数 θ 给定时，$X_1, X_2, \cdots, X_n$ 为相互独立随机样本，与 X 具有共同的分布 $F_X(x, \theta)$。由于风险的非齐次性，假设 θ 本身也是随机变量，具有分布函数 $\pi(\theta)$。这个分布在贝叶斯统计上称为先验分布，在精算学中称为结构分布。记

$$E\left(X_i|\theta\right) = \mu\left(\theta\right)$$

$$\mathrm{Var}\left(X_i|\theta\right) = \sigma^2\left(\theta\right)$$

$$E\left[\mu\left(\theta\right)\right] = \mu$$

$$E\left[\sigma^2\left(\theta\right)\right] = \sigma^2$$

$$\mathrm{Var}\left[\mu\left(\theta\right)\right] = \tau^2$$

现在要预测 (估计) 第 $n+1$ 年的损失 X_{n+1}，且将估计量限定在样本 $X_1, X_2, \cdots, X_n$ 的线性函数类

$$L = \left\{a_0 + \sum_{i=1}^{n} a_i X_i,\ a_i \in \mathbf{R}\right\}$$

中，并使得均方误差最小，即

$$\min_{\widehat{X_{n+1}} \in L} E\left[\left(X_{n+1} - \widehat{X_{n+1}}\right)^2\right] = \min_{a_i \in \mathbf{R}, i=0,1,\cdots,n} E\left[\left(X_{n+1} - a_0 - \sum_{i=1}^{n} a_i X_i\right)^2\right] \tag{2.5}$$

求解上面的最小化问题得到

$$\widehat{X_{n+1}} = Z \times \overline{X} + (1 - Z) \times \mu \tag{2.6}$$

这正是信度保费的形式，式中，$\overline{X}$ 为索赔平均，而 μ 为聚合保费，权重 $Z = \dfrac{n}{n+\kappa}$

为信度因子，并且

$$\kappa = \frac{E\left[\mathrm{Var}\left(X|\theta\right)\right]}{\mathrm{Var}\left(E\left(X|\theta\right)\right)} = \frac{\sigma^2}{\tau^2}$$

这里 σ^2 反映了数据内部之间的离散程度，而 τ^2 反映了风险参数 θ 的波动程度。

因此，由式 (2.6) 给出的信度保费是所有线性估计中具有均方预测误差最小的估计。事实上，从贝叶斯决策观点看，信度保费正是线性贝叶斯估计。根据贝叶斯定理，贝叶斯预测 $\widehat{X_{n+1}}^B = E\left(X_{n+1}|X_1, X_2, \cdots, X_n\right)$ 是所有样本可测函数中均方误差最小的估计，即

$$\widehat{X_{n+1}}^B = \underset{\widehat{X_{n+1}} \in G}{\arg\min} E\left[\left(X_{n+1} - \widehat{X_{n+1}}\right)^2\right] \tag{2.7}$$

式中，$G = \left\{g : g \text{ 是样本} X_1, X_2, \cdots, X_n \text{ 的可测函数}\right\}$。因为 $L \subset G$，所以

$$E\left[\left(X_{n+1} - \widehat{X_{n+1}}^C\right)^2\right] \geqslant E\left[\left(X_{n+1} - \widehat{X_{n+1}}^B\right)^2\right] \tag{2.8}$$

即贝叶斯预测比信度保费具有更小的均方预测误差。在信度理论中，称 $\widehat{X_{n+1}}^B$ 为贝叶斯保费。虽然贝叶斯保费比信度保费好，但是注意到，求解贝叶斯保费 $\widehat{X_{n+1}}^B$ 时需要样本 $X_1, X_2, \cdots, X_n$ 的分布以及风险参数 θ 的先验分布 $\pi(\theta)$ 的全部信息，而信度保费只需掌握 X_i 或 θ 的一二阶矩。不仅如此，在多合同信度模型中，这些矩都可以很容易由样本估计得到。

这些理论都发展于 20 世纪 60~70 年代，为最大精确度信度理论基本框架的构建奠定了基础。经过之后几十年的发展，许多学者在贝叶斯框架下将最大精确度信度理论推广到各个方向，信度理论的研究也逐步趋于成熟，在实际中得到广泛的应用。

2.3　信度理论的主要研究方向

Bühlmann (1967) 从贝叶斯观点出发，建立了无分布信度模型，得到了该模型下的信度保费公式，奠定了信度理论的统计基础。Bühlmann 和 Straub (1970) 从实际运用出发，引进保单索赔的自然权重，得到非齐次与齐次信度估计，并对结构参数提出了相应的估计，使得信度保费能直接运用于实际。这些结果已经成为信度理论的基本公式，称之为经典信度理论。后面的大部分关于信度理论的研究工作都是在这一理论框架下进行的。但是，随着精算科学在保险行业中的应用逐步精化、细化，保费价目表变得越来越精细，这些经典信度模型已经不能满足保险产品定价的需要。在逐步的研究过程中，学者们建立了适合实际需要的各种形式的信度模型。下面将目前信度理论的主要研究方向进行总结。

2.3.1 信度模型的拓展

Bühlmann (1967) 建立的信度模型可简单描述为独立同分布模型，即假设合同之间相互独立，个体样本索赔为条件独立同分布，且合同之间的风险参数也是独立同分布的。Bühlmann 和 Straub(1970) 虽然将自然权重引入保险合同，但是，仍然假设合同之间相互独立，样本有共同的条件期望，并且条件方差有特殊的结构。然而，在实际保险业务中，影响索赔的风险是非常复杂的，不仅在保单之间可能存在相依性，而且在时间分量上可能存在某些趋势。因此，根据实际需要，研究者对经典的 Bühlmann 信度模型、Bühlmann-Straub 模型进行了多方面的拓展。概括起来，有以下几个重要的拓展信度模型。

(1) Jwell 分层信度模型。从贝叶斯观点来看，分层信度模型类似于分层贝叶斯模型。在许多情况下，影响索赔的风险因素有多个，这时可根据这些因素逐个对保单进行细分类，得到各个因素下的保单子类，然后对每个子类进行信度保费定价。例如，汽车损害保险中，影响汽车索赔的因素有司机的性别、年龄、职业，车辆的类型、行驶区域等。首先根据性别因素将司机分为男性司机与女性司机，然后将不同性别的司机按年龄划分，再对具体的不同年龄、不同性别的司机按职业类型分类 …… 这样就得到多个因素对汽车保单的分类树形结构，在树形结构的每一层都可以计算其信度保费。其推广形式可参考 Sundt (1979)、Norberg (1986)、Bülmann 和 Jewell (1987) 等的研究。

(2) Hachemeister 回归信度模型。Hachemeister 在利用 Bülmann-Straub 信度模型对美国各州的汽车第三责任险进行定价时，注意到索赔数据在时间分量上由于通货膨胀的影响而具有时间趋势效应，提出用回归模型来刻画该效应，建立了回归信度模型，得到的保费估计为聚合保费与回归模型的最小二乘估计的加权平均，可参考 Hachemeister (1975) 的研究。随后，许多学者都在 Hachemeister 回归信度模型的基础上进行讨论，如 Bawelinckx 和 Goocaerts (1990)、Cossette 和 Luong (2003) 等。

(3) 多维信度模型。在某些情况下，需要估计的风险参数为包含多个分量的随机向量。用类似于经典信度理论的方法，将随机向量看成一个整体从而可以建立多维信度模型，得到的估计称为多维信度估计。多维信度模型首先由 Jewell (1973) 提出，随后被应用到其他各种信度模型中，可参考 Jewell (1973; 1974)、Bawelinckx 和 Goocaerts (1990) 等的研究。

(4) 交叉分类信度模型。在分层信度模型中，利用各因素对风险进行详细分类，得到各层的信度保费估计。然而，各因素之间可能存在交叉效应。Goulet (2001) 讨论了两水平交叉分量信度模型，也可参考 Kaas 等 (1996) 的研究。

(5) 纵向数据信度模型。在回归信度模型中，可能存在其他某些固定效应，在

统计中可用纵向数据模型来刻画。Frees 等 (1999) 讨论了纵向数据模型下的信度保费，并利用纵向数据的统计理论对信度模型进行了解释。

(6) 广义线性模型。对经典回归模型的另一个推广是广义线性模型。在指数族分布的假设下，通过某个联系函数，将非线性模型转化为回归模型。关于广义线性模型与信度保费的讨论，可参考 Nelder 和 Verrall (1997)、Antonio 和 Beirlant (2006) 等的研究。

(7) SUR 信度模型。假设每个保单合同可分成一些单位元，对每个单位元都有回归效应，这时可将这些单位元合并建立 SUR 模型。在信度理论框架下，Pitselis (2004) 讨论了 SUR 信度模型，得到该模型下的信度估计，并讨论了结构参数的估计问题。

2.3.2　关于精确信度条件的研究

贝叶斯保费是所有样本的可测函数中使得预测均方误差最小的估计，而信度保费是所有样本的线性函数中使得预测均方误差最小的估计。因此，在一般情况下，贝叶斯保费比信度保费在均方误差意义下好。然而，当样本的条件分布与风险参数的先验分布满足一定条件时，贝叶斯保费与信度保费等价，这时称信度保费为精确信度。Jewell 于 1974 年证明：样本的条件分布为指数族分布，并且风险参数的先验分布为相应的自然指数族分布时，满足精确信度条件，即信度保费与贝叶斯保费相等。Tweedie (1984) 等将这个结论推广为指数扩散族。对精确信度进行研究的还有 Goel (1982)、Gerber (1995) 等。

2.3.3　信度保费的统计性质研究

从统计观点来看，信度保费是未来年索赔的一个估计，因此在统计上必须满足一定的性质；从贝叶斯决策观点来看，信度保费为线性贝叶斯估计，在形式上为个体均值与聚合均值的加权平均，并且结构参数 (超参数) 可以由样本来估计，得到经验贝叶斯估计。因此，有许多学者对信度保费的相关性质进行研究。

(1) 估计的相合性。作为一个统计量，必须满足统计量的最重要的一个性质：相合性。Schmidt (1991) 提出信度估计的相合性问题，并证明了经典的信度理论中贝叶斯保费与信度保费都是风险保费 (个体保费) 的相合估计。Pan 等 (2008) 对 Esscher 保费原理下的信度估计的相合性进行了讨论。

(2) 估计的稳健性。保险公司的数据中偶尔会出现外来数据 (如超大索赔、输入错误等)，但有时又无法找到合适的方法将数据进行分离，这时一般要求给出的估计量具有稳健性。信度保费公式中包含样本均值统计量，而均值是非稳健统计量，因此信度估计也是不稳健估计。关于信度估计的稳健化方法可参考 Gisler (1980)、Kunsch (1992)、Gisler 和 Reinhard (1993) 和 Pitselis (2007) 等的研究。

(3) 经验贝叶斯渐近最优性。由于单合同信度保费估计中仍然存在结构参数，在多合同信度模型中，可以用样本对结构参数进行估计，将信度估计中的结构参数用其估计量代替，这时称该估计为经验贝叶斯估计。将结构参数代入后的经验贝叶斯信度估计是否能与单合同的信度估计非常接近，这就是经验贝叶斯渐近最优问题。Norberg (1980) 首先对经验贝叶斯信度估计的渐近最优性进行了讨论，Mashayekhi (2002) 将经验贝叶斯信度的渐近最优性推广到 Bühlmann-Straub 模型，并得到更弱的条件。

(4) 线性马尔科夫性。Witting (1988) 讨论了信度保费的线性马尔科夫性，并证明了满足线性马尔科夫性的充分必要条件。

2.3.4 相依风险信度模型

经典的信度理论假设风险之间是相互独立的，并且在时间分量上的索赔也是条件独立的，然而在实际中存在许多风险相依情形。例如，夫妻之间的寿命呈现正相关，同一次交通事故可以导致多次索赔，地域临近的房屋面临共同的火灾风险等。类似地，同一保单在不同时间的索赔也可能具有相依性。在时间或风险之间相依的情况下建立的信度模型统称为相依信度。事实上，分层信度是一种风险之间相依的信度模型，而回归信度模型中考虑了时间分量具有相依的情况。对信度模型的合同风险相依以及时间分量上索赔相依的讨论可参考 Yeo 和 Valdez (2006)、Wen 等 (2009)、Purcaru 和 Denuit (2002；2003)、Frees 和 Wang (2005) 等的研究。

2.3.5 非参数信度模型

信度保费中同时含有先验信息与样本信息。然而，Young 认为样本分布中的参数容易由样本进行估计，但先验分布本身或先验分布的超参数很难估计，她提出用非参数的方法对先验分布进行估计，由此得到半参数信度模型，可参考 Young (1997；1998a；1998b；2000)、Huang 等 (2003) 或 Qian (2000) 提出的用核密度估计对先验分布进行估计，并利用非参数方法讨论信度估计的相合性的研究。

2.3.6 信度理论中结构参数的估计

Bühlmann 和 Straub (1970) 根据矩方法对 Bühlmann-Straub 信度模型中的结构参数 μ、σ^2、τ^2 提出了相应的估计。然而，在统计中有很多方法可以估计参数，并能使得到的参数估计满足一些好的统计性质。因此，许多研究者从各个角度寻找更好的估计。Vylder (1978) 对回归模型中的结构参数矩阵与结构参数向量提出一些估计，并讨论这些估计的性质。他在 1981 年又提出结构参数估计的问题，在最小方差的意义下对参数估计的最优性进行了讨论。Norberg (1982) 也讨论了结构参数估计的最优性问题。Vylder 和 Goovaerts (1991) 在零峰度假设下讨论了结构参数估计的最优性问题。这方面的研究具体可参考 Vylder 和 Goovaerts

(1991；1992a；1992b；1992c)、Goulet (1998)、Norberg (1982)、Cossette 和 Luong (2003) 以及 Lo 等 (2006) 的文章。

2.3.7　保费原理下的信度估计

经典的信度理论建立了净保费的估计模型。然而，在实际运用中，净保费原理不能满足正的安全负荷性，保险公司将根据自身对风险的态度选取合适的保费原理。首先注意到这个问题的是 Gerber (1980)，他注意到，经典信度理论得到的是净保费，这是因为它使用了平方损失函数。因此，Gerber 及相关学者将平方损失函数修改为指数加权损失函数，用类似于经典信度理论的方法，将估计限定在样本的线性组合下得到了 Esscher 保费原理下的线性信度保费估计。Centeno (1989) 考虑了方差保费原理下的信度估计。进而，Heilmann (1989) 将指数加权损失函数推广为其他损失函数，并用决策理论的观点对信度保费进行了解释。类似地，Schmidt 和 Timpel (1995) 考虑了一般的加权损失函数情形。对 Esscher 保费原理下的信度估计进行讨论的还有 Pan 等 (2008)。

2.3.8　信度模型在其他方面的应用

在保险精算中，信度理论已经成为保险定价的基本理论，至今已经发展得较为成熟。事实上，信度理论已经成功运用到保险精算学的其他各个领域。下面仅列出信度理论在精算其他领域的应用的一部分，具体可参考相应的参考文献：① 准备金评估，可参考 Venter (1989)、Vylder (1982) 等的研究；② 随机过程预测，如 Norberg (1992) 等的研究；③ Kalman 滤波，可参考 Jong (1983) 等的研究；④ 死亡率或失效率评估，可参考 Hardy 和 Panjer (1998)、Nielsen 和 Sandqvist (2005) 等的研究；⑤ 风险度量，可参考 Siu 和 Yang (1999) 等的研究。

第二篇

信度保费的估计方法

第 3 章　Bayes 保费与信度保费

3.1　保 费 估 计

设 $X_1, X_2, \cdots, X_n, \cdots$ 为某保单在各年的索赔额。一般地，假设 $X_i (i = 1, 2, \cdots)$ 相互独立且服从共同的分布 $F_X(x, \theta)$，其中，θ 是风险参数，反映了保单合同的风险特征。对保险精算师而言，一个重要的任务就是为保单合同 X 制定合理的价格。

定义 3.1.1　设风险 X 有 n 年的索赔记录 $X_1, X_2, \cdots, X_n$，X_{n+1} 表示未来一年的索赔，若用 $X_1, X_2, \cdots, X_n$ 的某个函数 $g(X_1, X_2, \cdots, X_n)$ 来估计 (预测)X_{n+1}，即

$$\widehat{X_{n+1}} = g(X_1, X_2, \cdots, X_n) \tag{3.1}$$

则称 $g(X_1, X_2, \cdots, X_n)$ 为一个保费估计。

在实际运用中，常用的有以下几个保费估计。

1. 样本均值

假设 $X_1, X_2, \cdots, X_n, X_{n+1}$ 独立同分布于 $F_X(x, \theta)$，$E(X_i) = \mu$，则样本均值 $\overline{X}$ 为风险 X 的一个较好的保费估计，由强大数定律有

$$\overline{X} \xrightarrow{\text{a.s.}} \mu = E(X_{n+1})$$

2. 聚合保费估计

若 $E(X_{n+1}) = \mu$ 已知，则 μ 是 X_{n+1} 的一个保费估计，称之为聚合保费，注意聚合保费与样本容量 n 无关。

3. 风险保费

设在 θ 给定下，$X_1, X_2, \cdots, X_n$ 独立同分布于 $F_X(x, \theta)$，而 $\theta \sim \pi(\theta)$，若 $F_X(x, \theta)$ 与 $\pi(\theta)$ 均为已知，则

$$\mu(\theta) = E(X_{n+1}|\theta) \tag{3.2}$$

是 X_{n+1} 的一个保费估计。由于 $\mu(\theta)$ 与风险参数 θ 有关，称之为风险保费。注意到，在保险实际中，由于风险参数 θ 一般是未知的，因此 $\mu(\theta)$ 本身也是未知的，需要由样本来估计。

4. Bayes 保费

设在 θ 给定下，$X_1, X_2, \cdots, X_n$ 独立同分布于 $F_X(x, \theta)$，而 $\theta \sim \pi(\theta)$，则预测均值

$$\widehat{X_{n+1}}^B = E(X_{n+1}|X_1, X_2, \cdots, X_n) \tag{3.3}$$

是 X_{n+1} 的一个保费估计，称之为 Bayes 保费。

3.2 贝叶斯保费估计

定理 3.2.1 若在 θ 给定下，$X_1, X_2, \cdots$ 独立同分布于 $F_X(x,\theta)$，而 θ 为随机变量，且 $\theta \sim \pi(\theta)$。若取平方损失如下:

$$L(X,a) = (X-a)^2$$

则最小化问题

$$\min_{g\in\mathcal{L}} E\left[X_{n+1} - g(X_1, X_2, \cdots X_n)\right]^2 \tag{3.4}$$

的解为

$$\widehat{X_{n+1}}^B = E(X_{n+1}|X_1, X_2, \cdots, X_n)$$

其中，

$$\mathcal{L} = \{g(X_1, X_2, \cdots, X_n), g \text{ 为 } X_1, X_2, \cdots, X_n \text{ 的可测函数}\}$$

证明： 为方便，记 $\varPhi = E\left\{[X_{n+1} - g(X_1, X_2, \cdots, X_n)]^2\right\}$，由条件期望公式，有

$$\varPhi = E\left(E\left\{[X_{n+1} - g(X_1, X_2, \cdots, X_n)]^2 | X_1, X_2, \cdots, X_n\right\}\right) \tag{3.5}$$

要使 $\varPhi$ 最小化，只需

$$u = E\left\{[X_{n+1} - g(X_1, X_2, \cdots, X_n)]^2 | X_1, X_2, \cdots, X_n\right\}$$

达到最小即可。由于

$$\begin{aligned} u &= E\left\{\left[X_{n+1}^2 - 2X_{n+1}g(X_1, \cdots, X_n) + g^2(X_1, \cdots, X_n)\right] | X_1, \cdots, X_n\right\} \\ &= E\left(X_{n+1}^2 | X_1, \cdots, X_n\right) - 2g(X_1, \cdots, X_n)E(X_{n+1}|X_1, \cdots, X_n) + g^2(X_1, \cdots, X_n) \end{aligned}$$

因为 u 是关于 g 的凸函数，令

$$\frac{\partial u}{\partial g} = -2E(X_{n+1}|X_1, \cdots, X_n) + 2g(X_1, \cdots, X_n) = 0$$

则有

$$g(X_1, X_2, \cdots, X_n) = E(X_{n+1}|X_1, X_2, \cdots, X_n)$$

因此最优估计为

$$\widehat{X_{n+1}}^B = E(X_{n+1}|X_1, X_2, \cdots, X_n)$$

定理证完。

定理说明，Bayes 保费 $\widehat{X_{n+1}}^B$ 为平方损失函数下的最优保费估计。

注记 3.2.1　注意到

$$\begin{aligned}\widehat{X_{n+1}}^B &= E(X_{n+1}|X_1, X_2, \cdots, X_n)\\ &= E\left[E(X_{n+1}|X_1, X_2, \cdots, X_n, \theta)|X_1, X_2, \cdots, X_n\right]\\ &= E\left[E(X_{n+1}|\theta)|X_1, X_2, \cdots, X_n\right]\end{aligned}$$

第三个等号成立是因为在 θ 给定下，$X_1, X_2, \cdots, X_n$ 与 X_{n+1} 相互独立。若记

$$E(X_i|\theta) = \mu(\theta)$$

则 Bayes 保费也可以表达为

$$\widehat{X_{n+1}}^B = E[\mu(\theta)|X_1, X_2, \cdots, X_n] = \int \mu(\theta)\,\pi(\theta|X_1, X_2, \cdots, X_n)\,\mathrm{d}\theta$$

例 3.2.1　设 X_i 独立同分布于 Possion (θ)，θ 服从 Gamma (α, β)，且有

$$P(X_i = x) = \frac{\theta^x}{x!}\mathrm{e}^{-\theta}, \quad x = 1, 2, \cdots$$

$$\pi(\theta) = \frac{\beta^\alpha}{\Gamma(\alpha)}\theta^{\alpha-1}\mathrm{e}^{-\beta\theta}, \quad \theta > 0$$

求 Bayes 保费。

解　风险保费为

$$\mu(\theta) = E(X_i|\theta) = \theta$$

又因为 θ 的后验分布正比于样本联合分布和先验分布的乘积，即

$$\begin{aligned}\pi(\theta|X_1, X_2, \cdots, X_n) &\propto \left(\prod_{i=1}^n f(X_i|\theta)\right)\pi(\theta)\\ &\propto \theta^{\sum\limits_{i=1}^n X_i+\alpha-1}\mathrm{e}^{-(\beta+n)\theta}\end{aligned}$$

因此后验分布为

$$(\theta|X_1, X_2, \cdots, X_n) \sim \mathrm{Gamma}\left(\sum X_i + \alpha, \beta + n\right)$$

则容易得到 Bayes 保费为

$$\widehat{X_{n+1}}^B = E(\theta|X_1, X_2, \cdots, X_n) = \frac{\sum\limits_{i=1}^n X_i + \alpha}{\beta + n}$$

例 3.2.2 设 X 独立同分布于 $b(1,\theta)$，θ 服从 Beta (a,b)，且有

$$P(X_i=x)=\theta^x(1-\theta)^{1-x},\quad x=0,1$$

$$\pi\left(\theta\right)=\frac{1}{B\left(a,b\right)}\theta^{a-1}\left(1-\theta\right)^{b-1},\quad 0<\theta<1$$

求 Bayes 保费。

解 由于

$$\widehat{X_{n+1}}^B=E\left[\mu\left(\theta\right)|X_1,X_2,\cdots,X_n\right]$$

而风险保费

$$\mu\left(\theta\right)=E\left(X_i|\theta\right)=\theta$$

后验分布为

$$\begin{aligned}\pi(\theta|X_1,X_2,\cdots,X_n)&\propto\theta^{a-1}\left(1-\theta\right)^{b-1}\left(\prod_{i=1}^{n}\theta^{X_i}(1-\theta)^{1-X_i}\right)\\&\propto\theta^{\sum\limits_{i=1}^{n}X_i+a-1}\left(1-\theta\right)^{n-\sum\limits_{i=1}^{n}X_i+b-1}\end{aligned}$$

因此有

$$(\theta|X_1,X_2,\cdots,X_n)\sim\text{Beta}\left(\sum_{i=1}^{n}X_i+a,n-\sum_{i=1}^{n}X_i+b\right)$$

所以 Bayes 保费为

$$\widehat{X_{n+1}}^B=E(\theta|X_1,X_2,\cdots,X_n)=\frac{a+\sum\limits_{i=1}^{n}X_i}{n+a+b}$$

例 3.2.3 设 X_i 独立同分布于 $u(0,\theta)$，而 θ 服从 Gamma (α,β)，求 $\widehat{X_{n+1}}^B$。

解

$$\begin{aligned}\widehat{X_{n+1}}^B&=E\left[\mu\left(\theta\right)|X_1,X_2,\cdots,X_n\right]\\&=E\left(\frac{\theta}{2}|X_1,X_2,\cdots,X_n\right)\\&=\frac{1}{2}E\left(\theta|X_1,X_2,\cdots,X_n\right)\end{aligned}$$

由于

$$\pi(\theta|X_1,X_2,\cdots,X_n)\propto\theta^{\alpha-1}\mathrm{e}^{-\beta\theta}\left[\prod_{i=1}^{n}\frac{1}{\theta}I(0<X_i<\theta)\right]$$

即

$$\pi(\theta|X_1,X_2,\cdots,X_n)=\frac{\theta^{\alpha-1}\mathrm{e}^{-\beta\theta}\dfrac{1}{\theta^n}I[X_{(n)}<\theta]}{\displaystyle\int_{X_{(n)}}^{\infty}\theta^{\alpha-1-n}\mathrm{e}^{-\beta\theta}\mathrm{d}\theta}$$

得

$$\widehat{X_{n+1}}^{B}=\frac{1}{2}\frac{\displaystyle\int_{X_{(n)}}^{\infty}\theta^{\alpha-n}\mathrm{e}^{-\beta\theta}\mathrm{d}\theta}{\displaystyle\int_{X_{(n)}}^{\infty}\theta^{\alpha-1-n}\mathrm{e}^{-\beta\theta}\mathrm{d}\theta}$$

注意到在这种情况下无法得到 Bayes 保费的显示表达式，在实际运用中，必须用数值积分方法计算上述积分。

从上面的几个例子可知，为求 Bayes 保费 $\widehat{X_{n+1}}^{B}$，必须知道样本分布 $F_X(x,\theta)$ 及先验分布 $\pi(\theta)$ 的所有信息。但是，即使样本分布和先验分布都已知，有时 Bayes 保费 $\widehat{X_{n+1}}^{B}$ 也没有显示解 (事实上，在大多数情况下均无显示解)。

3.3　信 度 保 费

针对 Bayes 保费的缺点，Bühlmann(1967) 提出信度理论解决了这个问题。他的想法是将 X_{n+1} 的估计 (预测) 限定在 $X_1,X_2,\cdots,X_n$ 的线性函数中，记

$$L(X,1)=\left(a_0+\sum_{i=1}^{n}a_iX_i|a_i\in\mathbf{R}\right)$$

在平方损失函数下，类似于求解 Bayes 估计，解下面的最小化问题：

$$\min_{g\in L(X,1)}E\left[X_{n+1}-g(X_1,X_2,\cdots,X_n)\right]^2=\min_{a_i\in\mathbf{R},i=0,1,\cdots,n}E\left(X_{n+1}-a_0-\sum_{i=1}^{n}a_iX_i\right)^2 \tag{3.6}$$

得到 X_{n+1} 的估计，即为著名的信度估计。

定理 3.3.1　假设在 θ 给定下，$X_1,X_2,\cdots,X_n,\cdots$ 独立同分布于 $F(x,\theta)$ 且 θ 服从先验分布 $\pi(\theta)$。记

$$\mu(\theta)=E(X_i|\theta),\quad \sigma^2(\theta)=\mathrm{Var}(X_i|\theta)$$

$$E[\mu(\theta)]=\mu,\quad E\left[\sigma^2(\theta)\right]=\sigma^2,\quad \mathrm{Var}[\mu(\theta)]=\tau^2$$

则求解最小化问题 (3.6) 得到 X_{n+1} 的信度估计:

$$\widehat{X_{n+1}}^{C}=Z\overline{X}+(1-Z)\mu$$