信度

本章主要内容

一、　　引言
二、　　信度的含义
三、　　有限波动信度
四、　　贝叶斯方法
五、　　一致最精确信度

一、引言

例 假设某保险公司开发一新险种，保单组合由 10 位投保人构成。开始，由于没有任何理赔经验数据，只能先验地假定他们具有相似的风险水平。然后假定每一投保人每年至多引发一次理赔，且理赔额为 1。最初，根据同行业的损失水平，估计这一保单组合的保费为 0.2，我们称这种保费为先验保费，或集体保费。

这样的估计是否符合实际情形，需要经验数据来验证。为了搜集足够的理赔数据，保险公司连续追踪十年，采集的全部数据显示于下表。

请分析下表的数据，说明保费收取是否合理，该如何改进。

	投保人

年份	1	2	3	4	5	6	7	8	9	10
1									1	
2	1	1							1	
3	1		1						1	
4			1						1	
5					1	1			1	
6		1								
7	1	1		1	1					
8	1			1						
9	1									
10	1									
$\overline{S}_i$	0.6	0.3	0.2	0.2	0.2	0.1	0	0	0.7	0
S_i	0.23									

（1）总体平均理赔额为 23/100＝0.23。

（2）投保人 9 和 1 的理赔记录明显偏高，0.7 与 0.6 的比例足以认为这二人的风险水平要劣于集体的风险水平；

（3）投保人 7、8 和 10 无理赔记录，表明他们的风险水平又优于集体的风险水平。

■ 经验费率厘定

经验费率厘定就是非寿险精算中用于消除风险子集的非同质性而发展起来的一类方法。这些方法主要包括两大类：

一类是在保险年度开始前，根据被保险人最近几个保险年度的理赔经验确定下一个保险年度的续期保费。

另一类是在保险年度末，根据被保险人当年的理赔经验来调整他在当年已经交纳的保险费。

■　信度理论

信度理论就是研究如何合理利用先验信息和个体理赔经验来进行估计，预测及制定后验保费。后验保费估计可以下面公式来表示

$$后验估计值 = z \times 经验值 + (1 - z) \times 先验值$$

其中 z（$0 \leqslant z \leqslant 1$）称为信度因子，后验保费估计值称为信度保费。只有正确地选择信度因子 z，才能保证调整后的保险费接近于真实的风险水平。

当 $z = 1$，称理赔经验具有完全信度。

当 $z = 0$ 时，称为理赔经验没有信度。

当 $0 < z < 1$ 时，称为理赔经验具有部分信度。

第一节 有限波动信度

一、数学模型

设 X_j 表示某投保人在过去时期 j 的损失或理赔， X_j 也可以看作某风险同质的保单组合中第 j 份保单的损失，或者风险分级系统中第 j 类的损失。 $j=1,\cdots,n$ 。假设 $E(X_j)=\xi$ （未知）， $\operatorname{var}(X_j)=\sigma$ ， $j=1,\cdots,n$ 。

设：

- $\bar{X}=n^{-1}(X_1+\cdots+X_n)$ 表示投保人或保单组合的损失经验，包括理赔次数、理赔强度和总理赔额
- M 表示先验获得的 ξ 的估计值。

问题：如何确定ξ的值？

方法一：完全可信，忽略 M，直接使用经验数据，即令$\xi = \bar{X}$ 。

方法二：没有信度，忽略过去的经验数据，直接令ξ＝M。

方法三：部分可信，取 M 和 $\bar{X}$ 的加权值。

$$\xi = z \times \bar{X} + (1-z) \times M$$

完全可信标准的含义

有限波动信度模型中，首先需要确定当经验数据达到多大规模时，才可以给其赋予 100％的可信度，这个数据规模也被称作完全可信度标准。如果经验数据的规模达到或超过这个标准，则经验数据的信度因子 $Z=1$，否则，其信度因子将小于 1。小于 1 的信度因子被称作部分可信度。

例如,精算师要估计汽车保险索赔频率，假设获得 1000 次出险次数的数据规模可以认为是达到完全可信标准，那么如果实际观测到 2000 次出险次数,则认为 2000 次出险的经验数据是完全可信的。若假设 2000 次的出险次数才达到完全可信标准，那么观测到 1000 次出险次数的经验数据就具有部分信度。

二、 索赔频率完全可信标准

问题：观察到的保单组出险次数达到多少时，索赔频率的估计值才是完全可信的？

例如，假设某风险同质性的保单组每个风险单位的索赔频率服从参数为 3 的泊松分布，那么有 100 个风险单位的同质保单组每年实际的索赔次数将在 300 附近变化。当泊松分布的参数（均值）比较大时，可以用均值为 300，方差为 300 的正态分布近似。

若要估计保单的索赔次数超过 330 的概率。由于 $\sqrt{300}=17.32$，因此 330 次出险次数意味着超出均值 300 的次数达到标准差的 $30/17.32=1.73$ 倍，$\Phi(1.73)=0.9582$，也就是说有 4.18%的概率可以观察到出险次数超过 330 次。

类似的，出险次数低于 270 的概率也为 4.18%。换句话说，观察到的保单组出险次数在期望次数 10%以内波动的概率为 95.82%。

这时我们可以认为索赔频率的估计值在 4.18%的置信度下是完全可信的

推广到一般情况：

某个体风险的索赔频率的均值为 μ_f，标准差为 σ_f，现用有 n 个风险单位的同质性保单组的总索赔次数 N 来估计索赔频率。保单组合的期望索赔次数为 $n\mu_f$。索赔频率的观察值 $\frac{N}{n}$ 落在区间 $[\mu_f - r\mu_f, \mu_f + r\mu_f]$ 的概率为

$$
\begin{aligned}
p &= \Pr(\mu_f - r\mu_f \le \frac{N}{n} \le \mu_f + r\mu_f) \\
&= \Pr(\frac{-r\mu_f}{\sigma_f/\sqrt{n}} \le \frac{\frac{N}{n} - \mu_f}{\sigma_f/\sqrt{n}} \le \frac{r\mu_f}{\sigma_f/\sqrt{n}}) \\
&= \Pr(\frac{-r\mu_f}{\sigma_f/\sqrt{n}} \le Z \le \frac{r\mu_f}{\sigma_f/\sqrt{n}})
\end{aligned} \qquad (1)
$$

如果上述概率不小于$1-\alpha$，则认为频率的经验值是完全可信的。

若定义

$$
y_\alpha = \inf_y \left\{ P\left(\left| \frac{\bar{X}-\mu_f}{\sigma_f/\sqrt{n}} \right| \le y \right) \ge \alpha \right\} \qquad (3)
$$

由（1）可得

$$r\frac{\mu_f}{\sigma_f/\sqrt{n}} \geq y_\alpha \qquad (2)$$

当风险单位数 n 足够大时，$Z=\dfrac{(N/n)-\mu_f}{\sigma_f/\sqrt{n}}$ 近似服从标准正态分布，则可以确定 y_α 的值。由式(1)和式(3)得

$$\alpha = P(|Z| \leq y_\alpha) = \Phi(y_\alpha) - \Phi(-y_\alpha) = 2\Phi(y_\alpha) - 1$$

因此，$y_\alpha = \Phi^{-1}((1+\alpha)/2)$

- 特例：泊松分布情形

如果假设个体风险的期望索赔次数服从参数为λ的泊松分布，$\mu_f=\sigma_f^2=\lambda$，n 个同质性风险单位的保单组合期望索赔次数为$n\lambda$，则（2）式为

$$r\frac{\lambda}{\sqrt{\lambda}/\sqrt{n}}\geq y_{\alpha}$$

整理可得，

$$n\lambda\geq\left(\frac{y_{\alpha}}{r}\right)^2\triangleq n_0$$

（4）期望索赔次数完全可信标准

$$n \geq \frac{1}{\lambda}\left(\frac{y_{\alpha}}{r}\right)^{2} \quad (5)$$ 风险单位数完全可信标准

因此，如果假设个体风险的索赔频率服从参数为λ的泊松分布，且要求索赔频率的观察值落在区间$[\lambda - r\lambda, \lambda + r\lambda]$的概率不小于$1-\alpha$，则同质性保单组合索赔次数的期望值应该满足式（4）给定的条件，或保单组合的风险单位数应满足式（5）的条件。这两个条件就是索赔次数的完全可信度标准，说明当期望索赔次数足够大时，或同质性保单组合的风险单位数目足够多，个体风险的经验索赔频率就具有完全的可信度。

表 1　**泊松分布索赔次数的完全可信度标准(n_0)**

$1-\alpha$	r						
	10%	7.5%	5%	4%	3%	2%	1%
20%	164	292	657	1026	1825	4106	16424
10%	271	481	1082	1691	3006	6764	27055
5%	384	683	1537	2401	4268	9604	38415
4%	422	750	1687	2636	4687	10545	42179
3%	471	837	1884	2943	5233	11773	47093
2%	541	962	2165	3382	6013	13530	54119
1%	663	1180	2654	4147	7372	16587	66349

- 非泊松分布情形

如果个体风险的索赔次数并非服从泊松分布，而是服从贝努利分布、二项或负二项等其他分布， 由（2）式，$r\dfrac{\mu_f}{\sigma_f/\sqrt{n}} \geq y_\alpha$，类似地利用正态近似得到，索赔次数的完全可信度标准（表示为索赔次数的期望值）为

$$n\mu_f \geq \left(\frac{y_\alpha}{r}\right)^2 \cdot \frac{\sigma_f{}^2}{\mu_f} = n_0 \frac{\sigma_f{}^2}{\mu_f} \tag{6}$$

其中 μ_f 和 σ_f 分别为索赔频率的均值和标准差， $n_0 = \left(y_\alpha / r\right)^2$ 。对于泊松分布，其均值和方差相等，因此上式的后一项为 1。

例1：You are given:

(i) The number of claims has probability function:

$$p(x)=\binom{m}{x}q^{x}(1-q)^{m-x}, \qquad x=0,\ 1,\ 2,\ldots,\ m$$

(ii) The actual number of claims must be within 1% of the expected number of claims with probability 0.95.

(iii) The expected number of claims for full credibility is 34,574.

Determine q.

解：二项分布的均值$\mu_f=mq$，方差$\sigma_f^2=mq(1-q)$。根据（6），

$$n\mu_f \geq n_0 \frac{\sigma_f^{\ 2}}{\mu_f}\text{，}$$

$$n_0 = \left(y_{0.95} / r\right)^2 = \left(1.96 / 0.01\right)^2 = 38416$$

$$34574 = 38416 \frac{mq(1-q)}{mq}$$

即 $(1-q) = \dfrac{34574}{38416} = 0.9$， $q = 0.1$。

三、索赔强度的完全可信标准

索赔强度是指一次理赔事件中的期望理赔额。$X_1,\cdots,X_n$ 表示某风险同质性保单发生的 **n 次索赔金额**，其中 X_j 表示第 j 次索赔事件的索赔额，独立同分布。假设 $E(X_j)=\mu_X$，$Var(X_j)=\sigma_X^2$ (j=1, 2, …, n)，即认为每次索赔金额期望和方差都是稳定的。索赔强度可以用样本均值 $\bar{X}=\frac{1}{n}\sum_{i=1}^{n}X_i$ 来估计。容易知道，

$$E(\bar{X})=\mu_X,\ Var(\bar{X})=\sigma_X^2/n。$$

当索赔强度估计值$\bar{X}$位于真值μ_X的一个较小区间 100r%的概率很高(100p%)时，保险公司将对个体风险的索赔强度经验数据直接赋予完全可信性，即

$$P\left(-r\mu_X \le \bar{X}-\mu_X \le r\mu_X\right) \geqslant \alpha, \quad r>0, \quad 0<\alpha<1 \qquad (7)$$

由式(7)，有

$$P\left(\left|\frac{\bar{X}-\mu_X}{\sigma_X/\sqrt{n}}\right| \le \frac{r\mu_X\sqrt{n}}{\sigma_X}\right) \geqslant \alpha \qquad (8)$$

由中心极限定理知，当 n 充分大时，

$z=\frac{\bar{X}-\mu_X}{\sigma_X/\sqrt{n}} \to N(0,1)$，因此 $r\mu_X\sqrt{n}/\sigma_X \geqslant y_\alpha$，

整理可得索赔强度的完全可信标准

$$n \geqslant n_0 \frac{\sigma_X{}^2}{\mu_X{}^2} = n_0 CV_X^2, \quad (9)$$

式中，$n_0 = n_0 = (y_\alpha / r)^2$，$y_\alpha = \Phi^{-1}[(1+\alpha)/2]$，$CV_X$是索赔强度的变异系数。**式(9)表明：如果经验数据是完全可信的，则个体风险的索赔次数应大于 $n_0 CV_X^2$。**

当σ_X和μ_X的真实值未知时，可用样本标准差和样本均值来代替。

例 2：设 X 表示某个体保单发生一次索赔的索赔额，服从均值为 2 的指数分布，假定 $r=0.05$，$\alpha=0.9$，求索赔强度的完全可信标准。

【解】根据指数分布的性质，μ_X =2，σ_X^2 =4， 设 n 表示观察到该保单的索赔次数，则由（9）知

$$n \geqslant n_0 \frac{\sigma_X{}^2}{\mu_X{}^2} = 1082.41 \times \frac{4}{4} = 1082.41$$

因此，当观察到保单索赔次数 n 大于 1082 时，保单索赔额经验数据完全可信。

四、纯保费的完全可信标准

纯保费是指一个风险在单位时期内的期望总赔款，即保单组合总赔款与风险单位数之比。例如，有 50 辆汽车的保单组合在一年内发生了 25000 元的赔款，则纯保费的观察值为每车年 25000/50＝500 元。纯保费还可以表示为索赔频率与索赔强度的乘积。由于纯保费受到索赔次数和赔付额的共同影响，所以其波动性更大。对于给定的个体风险，由于随机波动的影响所造成的纯保费的方差被称作过程方差。

设 PP 表示单位时期内该风险的纯保费, N 表示索赔频率, Y_i 表示第 i 次索赔额, 假设每次索赔额 X_i 独立同分布, 且都与索赔次数 N 独立。则个体风险纯保费为:

$$PP=X_1+X_2+\cdots+X_N$$

由集体风险模型知

$$\mu_{PP} = \mu_f \mu_X$$

$$\sigma_{PP}^2 = \mu_f \sigma_X^2 + \mu_X^2 \sigma_f^2$$

设 $PP_1, \cdots, PP_n$ 表示 n 个风险单位同质性保单组纯保费的经验数据，其样本均值 $\bar{P}_n = \frac{1}{n}\sum_{i=1}^{n} PP_i$ 作为个体风险的纯保费观测值。

其落在期望纯保费μ_{PP}的一个较小区间 100r%的概率为

$$p = \Pr(\mu_{PP} - r\mu_{PP} \le \bar{P}_n \le \mu_{PP} + r\mu_{PP})$$
$$= \Pr(-r(\frac{\mu_{PP}}{\sigma_{PP}/\sqrt{n}}) \le \frac{\bar{P}_n - \mu_{PP}}{\sigma_{PP}/\sqrt{n}} \le r(\frac{\mu_{PP}}{\sigma_{PP}/\sqrt{n}}))$$

当保单组的索赔次数足够大，$\frac{\bar{P}_n - \mu_{PP}}{\sigma_{PP}/\sqrt{n}}$ 可以用正态分布来近似。因此，纯保费观测值在期望纯保费μ_{PP}的 r%以内波动的概率为α的完全可信标准（表示为风险单位数）为

$$n \ge n_0 \left(\frac{\sigma_{PP}}{\mu_{PP}} \right)^2 \quad （10）$$

其中 $n_0 = \left(y_\alpha / r \right)^2$， $y_\alpha = \Phi^{-1}((1+\alpha)/2)$ 。

如果进一步假设**索赔频率服从参数为**λ **的泊松分布**，则

$$\mu_{PP} = \lambda\mu_X$$

$$\sigma_{PP}^2 = \lambda(\mu_X^2 + \sigma_X^2)$$

将它们代入不等式（10），并经适当整理则有

$$n\lambda \geq n_0\left[1+\left(\frac{\sigma_X}{\mu_X}\right)^2\right] = n_0\left[1+CV_X^2\right] \quad (11)$$

$$n \geq \frac{n_0}{\lambda}\left[1+CV_X^2\right] \quad (12)$$

这就是纯保费的完全可信度标准（分别表示为期望索赔次数和风险单位数），其中（11）式右边的第一项是在泊松分布假设下索赔频率的完

全可信度标准，方括号内的第二项是**赔付额变异系数的平方**。而赔付额变异系数的平方乘以索赔频率的完全可信度标准，就是索赔强度的完全可信度标准，因此，纯保费的完全可信度标准等于索赔频率的完全可信度标准加上索赔强度的完全可信度标准。

小结

- 当**索赔频率为泊松分布**时，索赔频率、索赔强度和纯保费的完全可信标准。

表 2 索赔频率为泊松分布的完全可信标准。

	完全可信标准		
	索赔频率	索赔强度	纯保费
风险单位数（或保单期）	n_0/λ	$n_0\left(\dfrac{CV_X^2}{\lambda}\right)$	$\dfrac{n_0}{\lambda}\left(1+CV_X^2\right)$
期望索赔次数	n_0	$n_0 CV_X^2$	$n_0(1+CV_X^2)$

- 如果**索赔频率并不服从泊松分布**，则纯保费的完全可信度标准（表示为期望索赔次数和风险单位数）分别为

$$n\mu_f \ge n_0\left[\frac{\sigma_f^2}{\mu_f}+\left(\frac{\sigma_X}{\mu_X}\right)^2\right] \quad (13)$$

$$n \ge n_0\left[\frac{\sigma_f^2}{\mu_f^2}+\frac{1}{\mu_f}\left(\frac{\sigma_X}{\mu_X}\right)^2\right] \quad (14)$$

在此式(13)中，如果假设索赔频率服从泊松分布，则方括号中的第一项等于 1。如果赔付额是一个常数，其方差为零，则上式就是索赔频率的完全可信度标准。总之，式（13）和（14）概括了前述所有的完全可信度标准，它们都可以看作此式的特殊形式。

表 3 索赔频率为非泊松分布的完全可信标准。

	完全可信标准		
	索赔次数	索赔强度	纯保费
风险单位数（保单期）	$n_0\left(\dfrac{\sigma_f^2}{\mu_f^2}\right)$	$n_0\left(\dfrac{\sigma_X^2}{\mu_X^2\mu_f}\right)$	$n_0\left(\dfrac{\sigma_f^2}{\mu_f^2}+\dfrac{\sigma_X^2}{\mu_X^2\mu_f}\right)$
期望索赔次数	$n_0\left(\dfrac{\sigma_f^2}{\mu_f}\right)$	$n_0 CV_X^2$	$n_0\left(\dfrac{\sigma_f^2}{\mu_f}+\dfrac{\sigma_X^2}{\mu_X^2}\right)$

例 3

已知某公司汽车险保单的理赔额和理赔次数分布满足以下条件:

(1)对于每一个被保险人来说,理赔次数服从泊松分布,参数为 λ;

(2)理赔额服从参数为 $\alpha=6$, $\theta=0.5$ 的帕累托分布;

(3)理赔额和理赔次数是独立的;

(4)完全可信标准为在 0.02 倍均值范围内波动的概率为 0.9。

求达到纯保费完全可信标准的最少期望理赔次数。

【解】 $n_0=\left(y_\alpha/r\right)^2=\left(\frac{1.645}{0.02}\right)^2$， μ_X =0. 5/5=0. 1,

$\sigma_X^2=\frac{2\times(0.5)^2}{5\times(4)}-(0.1)^2=0.015$

代入纯保费完全可信标准（11）可知

$$\left(\frac{1.645}{0.02}\right)^2\left(1+\frac{0.015}{0.1^2}\right)=16912.66$$

因此，为使总理赔额估计值完全可信，总理赔次数最低不能少于16913。

例 4：已知商业汽车责任险的一下信息：

（1）索赔频率服从均值为λ的泊松分布，λ服从$\alpha=1.5$，$\theta=0.2$的伽马分布

（2）索赔大小相互独立且服从均值为 5000 的指数分布

（3）r=0.05，p=0.9

求达到纯保费完全可信标准的期望索赔次数。

解： 设 PP 表示总索赔额， $PP = X_1 + X_2 + \cdots + X_N$ ，利用混合泊松分布的性质知，N 服从负二项分布， $r = a = 1.5$ ， $\beta = \theta = 0.2$

$$\mu_f = r\beta = 0.3, \sigma_f^2 = r\beta(1+\beta) = 0.36$$

$$\mu_X = 5000, \sigma_X^2 = 25000000$$

$$E(PP) = 0.3(5000) = 1500$$

$$Var(PP) = 0.3(25000000) + 0.36(25000000) = 16500000$$

由（10）达到完全可信标准所需的同质风险单位数为

$$n = (1.645/0.05)^2(16500000/1500^2) = 7937.67$$

因此，达到完全可信标准所需的期望索赔次数

$$n\mu_f = 0.3(7937.67) = 2381$$

例 5

已知一个团体牙医保险计划，每个被保险人的看牙次数服从参数为（3，Q）二项分布， Q 的值随被保险人变化而变化，服从（0，1）上的均匀分布，索赔金额服从参数为 $\mu=1000,\theta=5$ 的逆高斯分布。求达到纯保费的完全可信标准（r=0.05，p=0.9）的期望索赔次数。

解： $\mu_f=E(N)=E(E(N\,|\,Q)=E(3Q)=3(0.5)=1.5$

$$\begin{aligned}\sigma_f^2 &= E\left[Var(F\mid Q)\right]+Var\left[E(F\mid Q)\right]\\ &= E\left[3Q(1-Q)\right]+Var(3Q)\\ &= 3(0.5)-3(0.5)+3^2\left(\tfrac{1}{12}\right)\\ &= 1.25\end{aligned}$$

$$\mu_X = 1000, \sigma_X^2 = \frac{1000^3}{5}, CV_X^2 = \frac{1000^3/5}{1000^2} = 200$$

$$n_F = 1082.41\left(\frac{\sigma_f^2}{\mu_f}+CV_X^2\right) = 1082.41\left(\frac{1.25}{1.5}+200\right) = 217384$$

例 6 suppose past losses $X_1,\ldots,X_n$ are available for a particular policyholder. The sample mean is to be used to estimate $\xi = E(X_j)$. Determine the standard for full credibility. Then suppose there were 10 observations with 6 being zero and the others being 253, 398, 439, and 756. Determine the full-credibility standard for this case with $r = 0.5$ and p=0.9.

解：根据投保人的索赔经验，我们可以计算出他的索赔额的期望和标准差无偏估计值分别为 184.6 和 267.89。

根据完全可信条件为

$$n \geq n_0 \frac{\sigma_X^2}{\mu_X^2},$$

得

$$n > 1082.41\left(\frac{267.89}{184.6}\right)^2 = 2279.51$$

因此，至少需要有 2279 个样本数才能保证这个投保人的索赔经验是完全可信的。

例 6 续 Suppose that past losses $X_1,\ldots,X_n$ are available for a particular policyholder and it is reasonable to assume that X_j are independent and compound Poisson distributed, that $X_i = Y_1^{(i)} + Y_2^{(i)} \cdots + Y_{N_i}^{(i)}$, where each N_i is Poisson with parameter λ and the claim size distribution Y has mean θ_Y and Variance σ_Y^2.

Determine the standard for full credibility when estimating the expected number of claims per policy and then when estimating the expected dollars of claim per policy.

Then Determine if these standards are met for the data in example 2, where it is now known that the first three nonzero payments came from a single claim but the final one was from two claims, one for 129 and the other for 627.

解： 设 $N_1,\ldots,N_n$ 表示该投保人的理赔次数，则为估计平均理赔次数的完全可信条件

$$n \ge n_0 \frac{\mu_f}{(\mu_f)^2} = 1082.41\left(\frac{\lambda}{\lambda^2}\right),$$

由于 10 年内总共有 5 次理赔，所以 λ 的估计值为 0.5，因此完全可信条件为

$$n \ge 1082.41\left(\frac{0.5}{0.5^2}\right) = 2164.82,$$

显然，投保人的 10 个观察数据量远远不满足完全可信条件。

五、 有限波动部分信度

当个体风险的经验数据规模没有达到完全可信标准，就可以给其赋予部分可信度。对于满足完全可信度标准的数据，对索赔频率、索赔强度和纯保费的估计值直接等于相应的样本均值，而对于部分可信度的数据，样本均值不具有完全信度，这时采用先验估计和样本均值的加权作为索赔频率、索赔强度和纯保费的估计值，即

$$z\bar{X}+(1-z)M \quad (15)$$

上述估计值称为信度估计值，z 称为信度因子，M 为投保人所在的风险子集得到先验估计,称为信度补项。

下面说明如何得到有限波动信度因子的值。

记 X_{full} 表示采用完全可信的经验数据得到的索赔频率估计值。

$X_{partial}$ 表示采用不完全可信的经验数据计算出的经验索赔频率。显然，$X_{partial}$ 和 X_{full} 的均值应该相等，不妨记作 μ 。

不妨用 $\sigma_{partial}$ 和 σ_{full} 分别表示 $X_{partial}$ 和 X_{full} 的标准差。由于计算 $X_{partial}$ 的样本量较小，所以，$\sigma_{partial} > \sigma_{full}$.

有限波动信度法认为，无论是采用部分可信度的数据对索赔频率进行估计，还是采用完全可信度的数据对索赔频率进行估计，这两个估计值应该具有相同的波动性。因此，z的选取应该使得

$$\Pr(\mu - r\mu \le X_{\text{full}} \le \mu + r\mu)$$
$$= \Pr(z\mu - r\mu \le zX_{\text{partial}} \le z\mu + r\mu)$$

$$\Pr(-r\mu / \sigma_{\text{full}} \le \left(X_{\text{full}} - \mu\right) / \sigma_{\text{full}} \le r\mu / \sigma_{\text{full}})$$
$$= \Pr(-r\mu / z\sigma_{\text{partial}} \le \left(zX_{\text{partial}} - z\mu\right) / z\sigma_{\text{partial}} \le r\mu / z\sigma_{\text{partial}})$$

采用正态近似，可得

$$r\mu / \sigma_{\text{ful}l} = r\mu / z\sigma_{\text{partial}}$$

$$z = \sigma_{\text{full}} / \sigma_{\text{partial}} \quad （16）$$

即，信度因子等于完全可信的经验估计的方差与部分可信的经验估计的方差之比。

1、索赔频率的信度因子

假设们要估计一个同质性风险类别的索赔频率μ，样本不满足完全可信度标准，样本中只包含 M 个风险单位。例如，一组同质性车险保单组有 M 个驾驶员，每位驾驶员的理赔次数都是独立同分布。我们要估计每位驾驶员的潜在出险次数，现在已经观察到每位驾驶员的出险次数为m_i，记$\hat{\mu}_M$为样本数据计算得到的索赔频率的观测值，即

$$\hat{\mu}_M = \sum_{i=1}^{M} m_i / M$$

若经验数据未达到完全可信标准，$\hat{\mu}_M$的方差为

$$\sigma^2{}_{\text{partial}} = \sigma_M^2 = Var(\sum_{i=1}^{M} m_i / M) = (1/M^2)\sum_{i=1}^{M} Var(m_i)$$

特例：泊松分布

假设每位驾驶员的理赔次数服从参数为λ的泊松分布，则

$$\sigma_M^2 = (1/M^2)\sum_{i=1}^{M}\lambda = M\lambda/M^2 = \lambda/M$$

对于一个有 M 个驾驶员的同质性保单组，没有达到完全可信标准，记保单组期望理赔次数为 n，则有 $M\lambda = n$，

$$\sigma^2{}_{\text{partial}} = \sigma_M^2 = (1/M^2)\sum_{i=1}^{M}\lambda = M\lambda/M^2 = \lambda^2/n$$

若经验数据达到完全可信标准，n_0为完全可信标准下的保单组期望理赔次数则$\sigma_M = \sigma_{\text{full}} = \lambda/\sqrt{n_0}$，代入信度因子的表达式（14）

$$z=\frac{\sigma_{\text{full}}}{\sigma_{\text{partial}}}=\frac{\mu/\sqrt{n_0}}{\mu/\sqrt{n}}=\sqrt{\frac{n}{n_0}} \quad (17)$$

（17）式说明泊松分布的索赔频率的部分可信度等于已知数据的（期望）索赔次数与完全可信度所要求的最小索赔次数（完全可信度标准）的平方根，因此（17）被称为平方根原则。

在计算部分可信度时，**最好使用期望索赔次数**。如果无法得到期望索赔次数，**也可以用实际观察到的索赔次数近似**。

2、索赔强度和总索赔额的部分信度因子的一般公式。

设个体风险的损失变量为 X。这里 X 可以表示为非泊松分布索赔频率、索赔强度和纯保费，记 X 的均值为 μ_X 和标准差为 σ_X。假设有 m 个风险单位同质性保单组的部分可信样本 $X_1,\cdots,X_m$ 来估计 μ_X。$X_{\text{partial}}=\bar{X}$ 是 μ_X 的经验估计，具有部分可信度，$\sigma_{\text{partial}}^2=\dfrac{\sigma_X^2}{m}$。要使 X_{partial} 具有完全可信的，根据（6）、（9） 和（10）可知，m 的值应小于完全可信标准要求的最小风险单位数 $n_F=n_0(\sigma_X^2/\mu_X^2)$，采用完全可信数据得到的经验估计 X_{full} 对应的方差为 $\sigma_{\text{full}}^2=\mu_X^2/n_0$。于是，由（16）有

$$z=\frac{\sigma_{\text{full}}}{\sigma_{\text{partial}}}=\frac{\sqrt{\mu_X^2/n_0}}{\sqrt{\sigma_X^2/m}}==\sqrt{\frac{m}{n_0(\sigma_X^2/\mu_X^2)}}=\sqrt{\frac{m}{n_F}} \quad (18)$$

其中，$n_F=n_0(\sigma_X^2/\mu_X^2)$是非泊松分布的完全可信标准下的最小风险单位数。若计算出 z 大于 1，说明完全可信条件满足，直接令 z=1。 因此，由有限波动信度法得到信度因子的公式为：

$$z=\min\left(\sqrt{\frac{m}{n_0\sigma_X^2/\mu_X^2}},1\right)=\min\left(\sqrt{\frac{m}{n_F}},1\right) \quad (19)$$

例 8

假设总索赔额 S 服从复合泊松分布,索赔次数的参数为 λ,索赔强度的均值为μ_X,方差为σ_X^2 ,设 $S_1, S_2, \cdots, S_m$是 m 个时期内的总索赔额观察值,求其信度因子。

【解】利用复合泊松分布的先验均值和方差公式计算得到

$$\mu_S = \lambda \mu_X, \quad \sigma_S^2 = \lambda(\mu_X^2 + \sigma_X^2)$$

完全可信标准的最小风险单位数是

$$n_F = n_0\left(\frac{\sigma_S^2}{\mu_S^2}\right) = n_0\left(\frac{\lambda(\mu_X^2 + \sigma_X^2)}{\lambda^2 \mu_X^2}\right) = \frac{n_0}{\lambda}\left(1 + \frac{\sigma_X^2}{\mu_X^2}\right)$$

代入式(19)得

$$z=\min\left(\sqrt{\frac{m}{\frac{n_0}{\lambda}\left(1+\sigma_X^2/\mu_X^2\right)}},\quad 1\right)=\min\left(\sqrt{\frac{m\lambda/n_0}{\left(1+\sigma_X^2/\mu_X^2\right)}},\quad 1\right) \quad (20)$$

例 9

某一般责任险保单组包含 2500 个被保险人, 已知下列信息:

(1) $X_i=\sum_{j=1}^{N_i} Y_{ij}$ 表示第 i 个被保险人在一年内发生的总理赔额;

(2) $N_1, N_2, \cdots, N_{2500}$ 独立同分布, 服从负二项分布, $r=2$, $\beta=0.2$;

(3) $Y_{i1}, Y_{i2}, \cdots, Y_{iN_i}$ 独立同分布，服从帕累托分布，参数

$\alpha=3.0$, $\theta=1000$;

(4)完全可信标准是在 5%的期望总理赔额范围内的概率为 90%。运用有限波动信度法求这个保单组年总理赔额的部分信度因子。

【解】根据负二项分布和帕累托分布的性质可知

$E(N)=r\beta=0.4$, $Var(N)=r\beta(1+\beta)=0.48$

$E(Y)=\theta/(\alpha-1)=500$

$Var(Y)=\theta^2\alpha/[(\alpha-1)^2(\alpha-2)]=750000$

因此

$E(X)=0.4\times500=200$

$$Var(X)=0.4\times 750000+0.48\times 500^2=420000$$

完全可信条件所需的最小风险单位数为:

$$n_F=\left(\frac{1.645}{0.05}\right)^2\times\frac{420000}{200^2}=11365$$

$$z=\sqrt{2500/11365}=0.47$$

例 10

已知总索赔额 X 服从复合泊松分布,其中个别索赔额变量服从指数分布。某投保团体过去 m 个时期的总索赔额为 $S_1, S_2, \cdots, S_m$ 相互独立同分布。假设索赔次数数据的信度因子为 0.8,求总索赔额数据 $S_1, S_2, \cdots, S_m$ 的信度因子。

【解】假设每个时期团体的理赔次数服从参数为 λ 的泊松分布,m

个时期的期望理赔次数为 n=$m\lambda$　由公式（17）信度因子等于

$$z=\sqrt{\frac{n}{n_0}}=\sqrt{\frac{m\lambda}{n_0}}=0.8$$

因此 mλ/n_0=0.64。

假设个体理赔额服从均值为 θ 的指数分布，$\mu_X^2=\sigma_X^2=\theta^2$，则 $\sigma_X^2/\mu_X^2=1$ 。由公式（20）

$$z=\min\left(\sqrt{\frac{m\lambda/n_0}{\left(1+\sigma_X^2/\mu_X^2\right)}},\quad 1\right)=\min\left(\sqrt{\frac{0.64}{(1+1)}},\quad 1\right)=0.566$$

例 11

27. You are given:

(i) X_{partial} = pure premium calculated from partially credible data

(ii) $\mu = E[X_{\text{partial}}]$

(iii) Fluctuations are limited to $\pm k\mu$ of the mean with probability P

(iv) Z = credibility factor

Determine which of the following is equal to P.

(A) $\Pr[\mu - k\mu \le X_{\text{partial}} \le \mu + k\mu]$

(B) $\Pr[\mathrm{Z}\mu - k \le \mathrm{Z}X_{\text{partial}} \le \mathrm{Z}\mu + k]$

(C) $\Pr[\mathrm{Z}\mu - \mu \le \mathrm{Z}X_{\text{partial}} \le \mathrm{Z}\mu + \mu]$

(D) $\Pr[1 - k \le \mathrm{Z}X_{\text{partial}} + (1 - Z)\mu \le 1 + k]$

(E) $\Pr[\mu - k\mu \le \mathrm{Z}X_{\text{partial}} + (1 - Z)\mu \le \mu + k\mu]$