Chapter (2): Credibility theory.

(1) Introduction:

Given observations $x_1, \dots, x_n$ for n years, the insurance company looks at revising the premium value, based on these observations.

This process is called Credibility theory.

→ classical credibility (limited fluctuation credibility).

→ Bayesian credibility.

→ Buhlmann's credibility.

(2) Classical Credibility:

Let $X_1, X_2, \ldots, X_n$ iid with

$$E(X) = \mu \; ; \; Var(X) = \sigma^2 .$$

which value do we have to take for a new premium for the ~~new~~ coming year ?

$\longrightarrow$ manual premium M.

$\longrightarrow$ $\bar{X}_n = \frac{\sum X_i}{n}$ (Full Credibility)

$\longrightarrow$ $Z\bar{X}_n + (1-\cancel{Z})M$; $0 \le Z \le 1$.

(Partial Credibility).

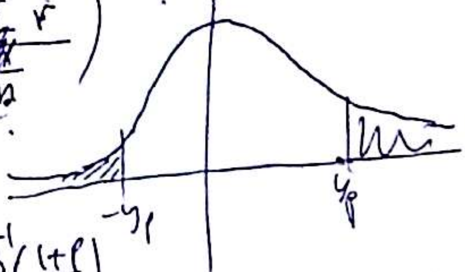
(2.1) Full Credibility:

- it means that we choose $\bar{X}_n$ as an approximative value of the premium. we should have that $\bar{X}_n$ is stable,
- which means that for r close to zero, and p close to 1, we have:

$$\mathbb{P}\left(\left| \frac{\bar{X}_n - \mu}{\mu} \right| \le r \right) \ge p .$$

- $\bar{X}_n \sim N\left(\mu ; \frac{\sigma^2}{n}\right)$, $\frac{\bar{X}_n - \mu}{\sigma/\sqrt{n}} \sim N(0,1)$

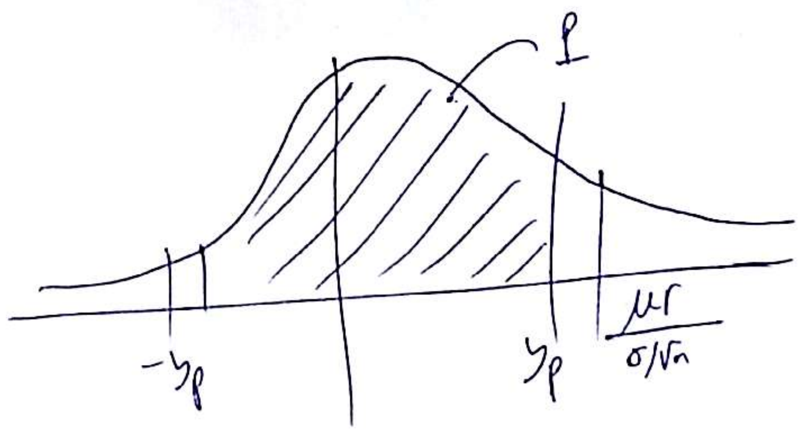
$$p \le \mathbb{P}\left(\left| \frac{\bar{X}_n - \mu}{\sigma/\sqrt{n}} \right| \le \frac{\mu r}{\sigma/\sqrt{n}} \right)$$



$Y \sim N(0,1)$ • Define y_p : $P(|Y| \le y_p) = p$

$$\phi(y_p) - \phi(-y_p) = p .$$

$$2\phi(y_p) - 1 = p \Rightarrow y_p = \phi^{-1}\left(\frac{1+p}{2}\right).$$



- $$\frac{\mu r}{\sigma/\sqrt{n}} \geq y_p = y$$

$$\sqrt{n} \geq \frac{y\sigma}{\mu r}.$$

$$n \geq \frac{y^2\sigma^2}{\mu^2 r^2}$$

$$\lambda_0 = \left(\frac{y}{r}\right)^2 \Rightarrow n \geq \lambda_0\left(\frac{\sigma}{\mu}\right)^2 = n_F.$$

- $n_F = \lambda_0\left(\frac{\sigma}{\mu}\right)^2$ is called standard of full credibility, measured in terms of the number of observations.

- in case we have a collective risk model, x_i = total loss in year i, N_i = number of claims in year i.

→ we define $x_F = n_F * E(X)$ the standard of full credibility, measured in terms of total loss.

→ we define $m_F = n_F * E(N)$, the standard of full credibility, measured in terms of total number of claims.

Example: let $N_1, \dots, N_n$, the number of Claims in the past n years. Let $X_1, \dots, X_n$ the tota loss in the past n years.

Assume each $X_i = \sum_{j=1}^{N_i} Y_{ij}$ with

$N_{ij} \sim$ Poisson (λ), $E(Y_{ij}) = \theta_Y$; $Var(Y_{ij}) = \sigma_Y^2$.

year 1

$Y_{11}, Y_{12}, \dots, Y_{1N_1}$	$Y_{21}, Y_{22} \dots Y_{2N_2}$	. . . . .
X_1	X_2	

Find the standard of full credibility, based on:

a) number of exposures units (observations).

b) total number of claims.

c) total amount of claims.

X is a Compound Poisson random variable with parameter λ. $X = \sum_{i=1}^{N} Y_i$.

$$\mu = E(X) = \lambda E(Y) = \lambda \theta_Y .$$

$$Var(X) = \lambda E(Y^2) = \lambda(\theta_Y^2 + \sigma_Y^2).$$

a) $$n_F = \lambda_0 \left(\frac{\sigma}{\mu}\right)^2 = \lambda_0 \frac{\lambda(\theta_Y^2 + \sigma_Y^2)}{\lambda^2 \theta_Y^2} = \frac{\lambda_0}{\lambda}\left(1 + \frac{\sigma_Y^2}{\theta_Y^2}\right).$$

b) $$m_F = n_F * E(N) = \lambda_0\left(1 + \frac{\sigma_Y^2}{\theta_Y^2}\right).$$

c) $$x_F = n_F * E(X) = \lambda_0 \theta_Y \left(1 + \frac{\sigma_Y^2}{\theta_Y^2}\right).$$

2.4 Partial Credibility:

We take:

$$P_c = Z.\bar{X}_n + (1-Z)M.$$

Z = Credibility factor.
P_c = Credibility premium.

In full Credibility Case, we have:

$$Var(\bar{X}_n) \leq \frac{\mu^2}{\lambda_0}.$$

So in Partial Credibility, we have:

$$\boxed{Var(P_c) \leq \frac{\mu^2}{\lambda_0}}$$

$$Var(P_c) = Var(Z\bar{X}_n) = Z^2 Var(\bar{X}_n)$$
$$= Z^2 \frac{\sigma^2}{n}.$$

$$Z^2 \frac{\sigma^2}{n} \leq \frac{\mu^2}{\lambda_0}.$$

$$Z^2 \leq \frac{\mu^2 n}{\sigma^2 \lambda_0} = \frac{n}{\frac{\sigma^2}{\mu^2}\lambda_0} = \frac{n}{n_F}.$$

$$Z \leq \sqrt{\frac{n}{n_F}}$$

$$Z = \inf\left(1, \sqrt{\frac{n}{n_F}}\right).$$

Example :

* $N \sim$ Poisson (10) : claim frequency
* $Y_{ij} \sim$ Pareto (3, 1) : claim size
* the full credibility standard is selected so that total claim (dollars) per exposure will be within 10% of expected total claim (dollars) per exposure 95% of the time. ($r = 0.1$; $p = 0.95$).

Find the Credibility factor based:

a) on 45 exposures (observations).

b) on a total claim number of 120.

c) on a total claim amount of 600.

---

a) $n_F = \frac{\lambda_0}{\lambda}\left(1 + \frac{\sigma_Y^2}{\theta_Y^2}\right) = ?$

$p = 0.95 \rightarrow \frac{1+p}{2} = \frac{1.95}{2} = 0.975 \rightarrow y = 1.96.$

$\lambda_0 = \left(\frac{y}{r}\right)^2 = \left(\frac{1.96}{0.1}\right)^2 = (19.6)^2 = 384.16$

$\rightarrow n_F = \frac{384.16}{10}\left(1 + \frac{0.75}{0.5^2}\right) = 124.85.$

$Z = \min\left(1, \sqrt{\frac{45}{124.85}}\right) = 0.6.$

b) $120 = E\left(\sum N_i\right) = n\,E(N) = 10n \Rightarrow n = 12.$

$Z = \min\left(1, \sqrt{\frac{12}{124.85}}\right) = 0.31.$

c) $600 = E\left(\sum_{i=1}^{n} X_i\right) = n\,E(X) = n\lambda\theta_Y = 5n.$

$\Rightarrow n = 120 \Rightarrow Z = \min\left(1, \sqrt{\frac{120}{124.85}}\right) = 0.98.$

Example: You are given:

* 350 claims with a total of 300,000 ($).
* manual premium of $M = 1000$.
* $z = 0.809$.

Determine the Credibility premium.

$$P_c = z\bar{X} + (1-z)M$$
$$= 0.809 * \frac{300,000}{350} + (1-0.809) * 1,000$$
$$= 884.42.$$

(3) Bayesian credibility:

- Suppose $X_1, \dots, X_n$ are n observations. We consider that risk is modelled through a parameter θ. We denote by $\pi(\theta)$, the prior distribution of θ.

$X_1, \dots, X_n \mid \Theta = \theta$ are iid.

- We call the Bayesian premium, the mean of the predictive distribution.

$$E(X_{n+1} \mid \underline{X} = \underline{x}) = \cancel{\int x_{n+1} f_{X_{n+1}|\underline{X}}(x_{n+1}|\underline{x})\, dx_{n+1}}$$

$$\cancel{= \int E(}$$

$$= \int x_{n+1} f_{X_{n+1}|\underline{X}}(x_{n+1} \mid \underline{x})\, dx_{n+1}.$$

$$= \int E(X_{n+1} \mid \Theta = \theta)\, \pi_{\Theta|\underline{X}}(\theta|\underline{x})\, d\theta.$$

- $E(X_{n+1} \mid \Theta = \theta)$ is called hypothetical mean.
- $E(X_{n+1})$ is called pure or Collective premium.

Ex: You are given the following for a dental insurer:

(i) claim number for an insured ~ Poisson.

→ (ii) $\frac{1}{2}$ of the insureds are expected to have 2 claims per year.

→ (iii) $\frac{1}{2}$ ———————, 4 claims per year.

An insured is chosen randomly and has made 4 claims in each of the first two policy years.

Determine the Bayesian estimate for this insured's claim number in the next (3^{rd}) policy year.

$X_1 = X_2 = 4$, $X_1, X_2 \sim \text{Poisson}(-)$

$$\theta = \begin{cases} \theta_1 & 2 \quad 1/2 \\ \theta_2 & 4 \quad 1/2 \end{cases}$$

$X_1, X_2 \mid \theta = 2 \sim \text{Poisson}(2)$.

$X_1, X_2 \mid \theta = 4 \sim \text{Poisson}(4)$

$X_1, X_2 \mid_{\theta=\theta} \sim \text{Poi}(\theta)$.

- model distribution:

$$f_{\underline{X}|\theta}(\underline{x}|\theta) = e^{-\theta}\frac{\theta^{x_1}}{x_1!} \cdot e^{-\theta}\frac{\theta^{x_2}}{x_2!} = \frac{1}{24^2} e^{-2\theta} \theta^{8}.$$

- Joint distribution

$$f_{\underline{X}\theta}(\underline{x},\theta) = \frac{1}{24^2} e^{-2\theta} \theta^8 \pi(\theta) = \frac{1}{1152} e^{-2\theta} \theta^8.$$

- marginal distribution:

$$f_{\underline{X}}(\underline{x}) = \frac{1}{1152} c \cdot \frac{\cancel{}}{\cancel{}} = \cancel{0.0114}.$$

- Posterior distribution: $c = e^{-4} 2^8 + e^{-8} 4^8$.

$\pi_{\theta|\underline{X}}(\theta|\underline{x}) = \frac{1}{c} \frac{\cancel{2^9}}{\cancel{\Gamma(9)}} e^{-2\theta} \theta^8 \cdot \cancel{\text{Gamma}(9, 2)}$.

- $E(X_3 | \underline{X} = \underline{x}) = \int \boxed{E(X_3 | \theta = \theta)} \pi_{\theta|\underline{x}}(\theta|\underline{x}) \, d\theta$

$$= \int \theta \, \pi_{\theta|\underline{x}}(\theta|\underline{x}) \, d\theta$$

$$= E(\theta|_{X=x}) = 2 \frac{1}{c} e^{-4} 2^8 + 4 \frac{1}{c} e^{-8} 4^8.$$

$$= \frac{2^9 e^{-4} + 4^9 e^{-8}}{2^8 e^{-4} + 4^8 e^{-8}} = 3.65.$$

Ex: $\theta \sim \text{Gamma}(5; 0.0005)$ 2000

Claim amount $X|_{\theta=\theta} \sim \text{Exp}(\theta)$.

$x_1 = 2000$; $x_2 = 1000$; $x_3 = 3{,}000$.

a) Compute the Bayesian premium without using hypothetical means?

b) ______ using hypothetical means?

a) $E(X_4 | \underline{X} = \underline{x}) = \int x \, f_{X_4|\underline{X}}(x|\underline{x}) \, dx$.

$$\boxed{f_{X|Y}(x|y) = \int f_{X|Z}(x|z) f_{Z|Y}(z|y) \, dz}$$

- model distribution: $f_{\underline{X}|\theta}(\underline{x}|\theta) = \theta e^{-\theta x_1} \theta e^{-\theta x_2} \theta e^{-\theta x_3} = \theta^3 e^{-6000\theta}$.

- Joint: $f_{\underline{X}\theta}(\underline{x}, \theta) = c \, \theta^3 e^{-6000\theta} \theta^4 e^{-2000\theta}$

$$= c \, \theta^7 e^{-8000\theta}$$

9

- Posterior: $\pi_{\theta|\underline{X}}(\theta|\underline{x}) \sim \text{Gamma}(8, 8000)$.

$$E(X_4 \mid \underline{X} = \underline{x}) = \int E(X_4 \mid \Theta = \theta)\, \pi_{\Theta \mid \underline{X}}(\theta \mid \underline{x})\, d\theta$$

$$= \int \frac{1}{\theta}\, \pi_{\Theta \mid \underline{X}}(\theta \mid \underline{x})\, d\theta$$

$$= E\left(\frac{1}{\Theta} \,\middle|\, \underline{X} = \underline{x}\right) =$$

$$= \int \left(\frac{\Gamma(8)}{8000^8}\right) \frac{1}{\theta}\, \theta^7 e^{-8000\theta}\, d\theta$$

$$= \left(\frac{\Gamma(8)}{8000^8}\right) \boxed{\int \theta^6 e^{-8000\theta}\, d\theta}$$

$$= \frac{\Gamma(8)}{8000^8} \cdot \frac{8000^7}{\Gamma(7)} = \frac{7}{8000} =$$

$$= \boxed{\frac{8000}{7}} = 1142.12.$$

4 Buhlmann' Credibility:

4.1 Credibility Premium:

$X_1, \dots, X_n$ random Variables.

let Y another random Variable.

We want to find $\alpha_0, \alpha_1, \dots, \alpha_n$ such that

$E\left(Y - (\alpha_0 + \alpha_1 X_1 + \dots + \alpha_n X_n)\right)^2$ is the

minimum.

let $g(\alpha_0, \alpha_1, \dots, \alpha_n) = E\left(Y - (\alpha_0 + \alpha_1 X_1 + \dots + \alpha_n X_n)\right)^2$

$$\frac{\partial g}{\partial \alpha_0} = E\, 2\left(Y - (\alpha_0 + \alpha_1 X_1 + \dots + \alpha_n X_n)\right)(-1) = 0$$

$$E(Y) = \alpha_0 + \alpha_1 E(X_1) + \dots + \alpha_n E(X_n).$$

$i = 1 \dots n$
$$\frac{\partial g}{\partial \alpha_i} = E\, 2\left(Y - (\alpha_0 + \alpha_1 X_1 + \dots + \alpha_n X_n)\right)(-X_i) = 0.$$

$$E(YX_i) = \alpha_0 E(X_i) + \alpha_1 E(X_1 X_i) + \dots + \alpha_n E(X_n X_i).$$

$$E(Y)E(X_i) = \alpha_0 E(X_i) + \alpha_1 E(X_1)E(X_i) + \dots + \alpha_n E(X_n)E(X_i)$$

Normal equations
$$\begin{cases} Cov(Y, X_i) = \alpha_1 Cov(X_1, X_i) + \dots + \alpha_n Cov(X_n, X_i) \\ E(Y) = \alpha_0 + \alpha_1 E(X_1) + \dots + \alpha_n E(X_n) \end{cases}$$

Example: let $X_1, \dots, X_{20}$, $Y = X_{21}$ with
$E(X_i) = 2$, $Var(X_i) = 3$; $Cov(X_i, X_j) = 1.5$, $i \neq j$.
Determine $\hat{\alpha}_0, \hat{\alpha}_1, \dots, \hat{\alpha}_{20}$?

$$\begin{cases} 2 = \alpha_0 + 2\alpha_1 + \cdots + 2\alpha_n \\ 1.5 = 1.5 \sum_{j \neq i} \alpha_j + 3\alpha_i \end{cases}$$

$$\begin{cases} \alpha_0 + 2\sum_{j=1}^{20} \alpha_j = 2. \\ 1.5 \sum_{j=1}^{20} \alpha_j + 1.5\alpha_i = 1.5 \Rightarrow \sum \alpha_j + \alpha_i = 1. \end{cases}$$

$$\alpha_0 = 2 - 2\sum_j \alpha_j = 2 - 2(1 - \alpha_i) = 2\alpha_i.$$

$$\alpha_0 + \sum_j^{20} \alpha_0 = 2 \rightarrow \alpha_0 = \frac{2}{21}.$$

$$\alpha_i = \frac{1}{21}$$

(4.2) Buhlmann's model:

* $X_1, \ldots, X_{n+1} | \theta$ independent.
* hypothetical mean $\mu(\theta) = E(X_i | \Theta = \theta)$.
  $\mu = E\mu(\Theta)$, $v = Var(\mu(\Theta))$,
* Process variance $v(\theta) = Var(X_i | \Theta = \theta)$.
  $a = E\,v(\Theta)$.
* Suppose now a new observation X_{n+1}.

The Buhlmann's Credibility premium for X_{n+1} is $\hat{\alpha}_0 + \hat{\alpha}_1 X_1 + \cdots + \hat{\alpha}_n X_n$, where $(\hat{\alpha}_0, \hat{\alpha}_1, \ldots, \hat{\alpha}_n)$ minimizes the function:

$$E\left(X_{n+1} - (\alpha_0 + \alpha_1 X_1 + \cdots + \alpha_n X_n)\right)^2.$$

- $E(X_1) = E\,E(X_1|\theta) = E\mu(\theta) = \mu$
- $E(X_1X_2) = E\,E(X_1X_2|\theta)$
$$= E\,E(X_1|\theta)E(X_2|\theta)$$
$$= E\,\mu^2(\theta)$$
$$= Var(\mu(\theta)) + (E\mu(\theta))^2$$
$$= v + \mu^2.$$
- $Cov(X_1, X_2) = v$
- $E(X_1^2) = E\,E(X_1^2|\theta)$
$$= E\left(var(X_1|\theta) + (E(X_1|\theta))^2\right)$$
$$= a + v + \mu^2.$$
- $var(X_1) = a + v$

$$\begin{cases} \alpha_0 + \mu \sum \alpha_j = \mu. \\ \sum_{j \neq i} \alpha_j v + \alpha_i (v + a) = v. \end{cases}$$

$$v \sum \alpha_j + a\alpha_i = v$$

$$\alpha_0 = \mu - \mu \sum \alpha_j = \mu - \frac{\mu}{v}(v - a\alpha_i) \qquad \mu\alpha_i = \frac{v}{a}\alpha_0$$
$$= \frac{\mu}{v} a \alpha_i \longrightarrow$$

$$\alpha_0 + \frac{a}{v} n = \mu, \quad \alpha_0 = \mu - \frac{a}{v} n.$$

$$\alpha_i = \frac{v}{a\mu}\alpha_0 = \frac{v}{a\mu}\left(\mu - \frac{a}{v}n\right)$$
$$= \frac{v}{a} - \frac{n}{\mu}.$$

- Credibility premium:

$$P_c = \hat{\alpha}_0 + \hat{\alpha}_1 X_1 + \dots + \hat{\alpha}_n X_n = \left(\mu - \frac{a}{v}n\right) + \left(\frac{v}{a} - \frac{n}{\mu}\right) n \bar{X}$$

$$\alpha_0 + \sum \frac{v}{a}\alpha_0 = \mu .$$

$$\alpha_0 \left(1 + \frac{vn}{a}\right) = \mu .$$

$$\alpha_0 = \frac{\mu}{1+\frac{Vn}{a}} = \frac{\mu a}{a+Vn}$$

$$\alpha_i = \frac{V}{a}\frac{\alpha_0}{\mu} = \frac{V}{a}\frac{a}{a+Vn} = \frac{V}{a+Vn} .$$

- Credibility Premium :

$$P_c = \hat{\alpha}_0 + \hat{\alpha}_1 X_1 + \cdots + \hat{\alpha}_n X_n .$$

$$= \frac{\mu a}{a+Vn} + \frac{V}{a+Vn} n\overline{X} .$$

$$= \frac{a}{a+Vn}\mu + \frac{Vn}{a+Vn}\overline{X} .$$

$$Z = \frac{Vn}{a+Vn} : \quad P_c = Z\overline{X} + (1-Z)\mu .$$

$$Z = \frac{n}{\frac{a}{V}+n} \quad ; \quad k = \frac{a}{V} ,$$

$$= \frac{n}{k+n} .$$

- $\hat{\alpha}_0 + \hat{\alpha}_1 X_1 + \cdots + \hat{\alpha}_n X_n$ is the best linear estimator of X_{n+1}, $E(X_{n+1}|\theta)$ and $E(X_{n+1}|\underline{X})$.

Scanned by CamScanner

Example: $X_1, \dots, X_n \mid \Theta = \theta \sim Exp(\theta)$,

$\Theta \sim Gamma(\alpha, \beta)$.

Determine the Buhlmann Credibility Premium for X_{n+1} (Buhlmann Credibility estimate for X_{n+1} / or $E(X_{n+1} \mid \theta)$ or $E(X_{n+1} \mid \underline{X})$).

$$P_C = Z\bar{X} + (1-Z)\mu.$$

$$Z = \frac{n}{k+n}; \quad k = \frac{a}{v}.$$

$$\mu = E\mu(\theta) = E\,E(X_1 \mid \theta) = E\frac{1}{\theta}$$

$$= \int \frac{1}{\theta} \frac{\beta^\alpha}{\Gamma(\alpha)} \theta^{\alpha-1} e^{-\beta\theta}\, d\theta.$$

$$= \frac{\beta^\alpha}{\Gamma(\alpha)} \int \theta^{(\alpha-1)-1} e^{-\beta\theta}\, d\theta.$$

$$= \frac{\beta^\alpha}{\Gamma(\alpha)} \frac{\Gamma(\alpha-1)}{\beta^{\alpha-1}} = \frac{\Gamma(\alpha-1)}{\Gamma(\alpha)}\beta.$$

α integer $\longrightarrow$ $\mu = \dfrac{\beta}{\alpha-1}$.

$$v = Var(E(X_1 \mid \theta)) = Var\left(\frac{1}{\theta}\right)$$

$$E\frac{1}{\theta^m} = \frac{\beta^\alpha}{\Gamma(\alpha)} \frac{\Gamma(\alpha-m)}{\beta^{\alpha-m}} = \beta^m \frac{1}{(\alpha-m)(\alpha-m+1)\cdots(\alpha-1)}.$$

$m=2$: $\quad E\dfrac{1}{\theta^2} = \dfrac{\beta^2}{(\alpha-2)(\alpha-1)}$.

$m=4$ $\quad E\dfrac{1}{\theta^4} = \dfrac{\beta^4}{(\alpha-4)(\alpha-3)(\alpha-2)(\alpha-1)}$.

$$a = E\,Var(X_1 \mid \theta) = E\frac{1}{\theta^2} = \text{—}$$

Ex: $X_1, \dots, X_n \mid \theta = n \sim Bin(n, p)$.

$$\theta = \begin{cases} 10 & 1/2 \\ 20 & 1/2 \end{cases}$$

Find ~~take~~ Buhlmann best estimate for X_{n+1}.

* $\mu = E\,\mu(\theta) = E\,p\theta = 15p$.
* $v = Var(\mu(\theta)) = Var(p\theta) = p^2\, Var(\theta)$
  $= p^2(-15^2 + 250) =$
* $a = E\, v(\theta) = E\, pq\theta = 15pq$.

$$X_{n+1} \sim \frac{n}{k+n}\bar{X} + \frac{k}{k+n}\mu.$$

Ex: $X_i \mid \theta$; θ : X_i = number of claims.

θ	0.05	0.3
$f_\theta(\theta)$	0.8	0.2

$x \mid \theta$	0	1	2
$f_{X\mid\theta}(x\mid\theta)$	2θ	θ	$1-3\theta$

Two claims are observed in year 1, calculate the ~~number~~ Buhlmann Credibility estimate for the number of claims in year 2.

* $\mu(\theta) = 0 + \theta + 2(1-3\theta) = 2 - 5\theta$.
* $v(\theta) = 0 + \theta + 4(1-3\theta) - (2-5\theta)^2$

$E(\theta) = 0.1$
$Var(\theta) = 0.01$

* $\mu = 2 - 5E(\theta) = 2 - 5(0.1) = 1.5$.
  $v = Var(2-5\theta) = 25\, Var(\theta) = 0.25$.
  $a = E(\theta + 4 - 12\theta - 4 - 25\theta^2 + 20\theta)$
  $= E(9\theta - 25\theta^2) = ~~26(0.01)~~ = 9(0.1) - 25(0.01 + 0.01)$
  $= ~~0.41~~\ 0.4$

$X_1 = 2$, $k = \frac{0.4}{0.25} = 1.6$

$$X_2 \sim \frac{1}{\frac{a}{v}+1}X_1 + \left(1 - \frac{1}{\frac{a}{v}+1}\right) \times (1.5) = 1.69.$$

Ex $X|\theta=\theta$ has mean θ and Var 500.

θ has mean 1,000 and Var 50

$X_1 = 750$, $X_2 = 1075$; $X_3 = 1{,}000$,

Buhlmann estimate for X_4?

* $\mu(\theta) = \theta \to \mu = 1{,}000$, $v = 50$

$v(\theta) = 500 \longrightarrow a = 500$

$k = \frac{500}{50} = 10$, $Z = \frac{3}{10+3} = \frac{3}{13}$

$X_4 \approx \frac{3}{13} \frac{750+1075+1{,}000}{3} + \frac{10}{13} \times 1{,}000$

$= 294.2 + 769.2 = 1063.4$.

4.3 Buhlmann - Straub model:

$X_1, \dots, X_n, X_{n+1} | \theta$ are independent.

X_i is the average of losses in year i:

$X_i = \frac{\sum_{j=1}^{m_i} X_{ij}}{m_i}$. given θ.

$\mu(\theta) = E(X_{ij}|\theta)$, $v(\theta) = Var(X_{ij}|\theta)$.

$\mu = E\mu(\theta)$; $v = Var(\mu(\theta))$; $a = E\,v(\theta)$.

- $E(X_1) = E X_{11} = E\,E(X_{11}|\theta) = E\mu(\theta) = \mu$.
- $Var(X_1) = \frac{Var(X_{11})}{m_1} = \frac{a}{m_1} + v$.
- $Cov(X_1, X_2) = E(X_1X_2) - E(X_1)E(X_2)$
  $= v$.
- $Var(X_1) = Var(E(X_1|\theta)) + E\,Var(X_1|\theta)$.
  $= v + \frac{E\,Var(X_{11}|\theta)}{m_1} = v + \frac{a}{m_1}$.

- $$\begin{cases} \mu = \alpha_0 + \alpha_1 \mu + \cdots + \alpha_n \mu = \mu \sum \alpha_j + \alpha_0 \\ v = \sum_{j=1}^{n} \alpha_j \operatorname{cov}(X_j, X_i) \end{cases}$$

$$= \sum_{j \neq i} \alpha_j v + \alpha_i \left(v + \frac{a}{m_i}\right)$$

$$= \sum \alpha_j v + \alpha_i \frac{a}{m_i}.$$

$$\Rightarrow v = v\left(1 - \frac{\alpha_0}{\mu}\right) + \alpha_i \frac{a}{m_i},$$

$$\alpha_i = v \frac{\alpha_0}{\mu} \frac{m_i}{a}.$$

$$\Rightarrow \mu = \alpha_0 + \frac{v \alpha_0}{a} \sum_{j=1}^{n} m_j, \quad m = \sum m_j$$

$$\hat{\alpha}_0 = \frac{\mu}{1 + \frac{vm}{a}}.$$

$$\hat{\alpha}_i = v \frac{m_i}{a} \cdot \frac{1}{1 + \frac{vm}{a}}.$$

- Credibility Premium:

$$\hat{\alpha}_0 + \sum_{j=1}^{n} \hat{\alpha}_j X_j = Z \bar{X} + (1 - Z)\mu.$$

$$\bar{X} = \frac{m_1 X_1 + \cdots + m_n X_n}{m},$$

$$Z = \frac{m}{m+k}, \quad k = \frac{a}{v}.$$

Example:
- In year j, there are N_j claims for m_j policies.
- Number of claims for one policy $\sim$ Poisson (λ),
- $\Lambda \sim$ Gamma (α, β).

Find ~~of~~ the Buhlmann–Straub estimate for the number of claims for one policy in year $n+1$.

* $\mu(\Lambda) = \Lambda$; $V(\Lambda) = \Lambda$.

$\mu = \alpha/\beta$; $v = \alpha/\beta^2$; $a = \alpha/\beta$.

$P_c = Z\bar{X} + (1-Z)\mu$, $Z = \frac{m}{m+k}$; $k = \frac{a}{v} = \beta$.

$= \frac{m}{m+\beta}\bar{X} + \frac{\beta}{m+\beta}\frac{\alpha}{\beta}$.

$= \frac{m}{m+\beta}\bar{X} + \frac{\alpha}{m+\beta}$.

Example :
- Number of clais per month ~ Poisson (λ).
- Λ ~ Gamma (~~5~~6, 100).
- Data:

month	Number of insured	Nber of clains	
1	100	6	600
2	150	8	
3	200	11	
4	300	?	
	m = 450		

$\bar{X} = \cancel{48.44} = \cancel{8.89} = 0.056$

$X_4 \simeq \frac{450}{450+100}(\cancel{8.89}\ 0.056) + \frac{6}{450+100}$.

$= \cancel{7.28}\ 0.057$

$0.057 \cancel{7.28} \times 300 = \cancel{2184} = \boxed{17.1}$

4.4 Exact Credibility :

Exact Credibility is when premium, calculated by Bayesian credibility is equal to premium calculated by Buhlmann model.

Some examples are given below.

$\pi(\theta)$	$f_{X\|\theta}(x\|\theta)$	$\pi(\theta\|\underline{x})$
Gamma	Poisson	Gamma
Normal	Normal	Normal
Beta	Bernoulli	Beta
Inverse gamma	Exponential	Inverse-gamma.

Example: $X|\theta=\theta \sim$ Poisson (θ)

$\theta \sim \text{Gamma}(\alpha, \beta)$.

$x_1, \ldots, x_n$.

$$\pi(\theta|\underline{x}) \sim e^{-\theta} \frac{\theta^{x_1}}{x_1!} \cdots e^{-\theta} \frac{\theta^{x_n}}{x_n!} \theta^{\alpha-1} e^{-\beta\theta}$$

$$\sim e^{-\theta(n+\beta)} \theta^{\sum x_i + \alpha - 1} \sim \text{Gamma}\left(\alpha + \sum x_i, n+\beta\right).$$

$$E(X_{n+1}|\underline{X}) = \int \theta \, \pi(\theta|\underline{X}) \, d\theta = \frac{\alpha + \sum x_i}{n+\beta}.$$

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

$$= \frac{n}{n+\beta}\bar{X} + \frac{\beta}{n+\beta}\frac{\alpha}{\beta}.$$

(5) Non-parametric Empirical Bayes estimation for the Buhlmann model.

- X random variable: $x_1, \ldots, x_n$.
  the estimate of the mean of X is $\bar{X}$.
  the estimate of the variance of X is
  $$s^2 = \frac{\sum_i (X_i - \bar{X})^2}{n-1}.$$
- In year i, we have $X_{i1}, \ldots, X_{ir}$ policies.

year \ policy	1	...	r
1	X_{11}	...	X_{1r}
⋮	⋮		⋮
n	X_{n1}	...	X_{nr}
average	$\bar{X}_1$	...	$\bar{X}_r$
variance	s_1^2	...	s_r^2

$$\hat{\mu} = \frac{\sum_{i=1}^{r} \bar{X}_i}{r} = \bar{X}.$$

$$\hat{a} = \frac{\sum_{i=1}^{r} s_i^2}{r}.$$

$$\hat{v} = \frac{1}{r-1} \sum_{i=1}^{r} \left(\bar{X}_i - \bar{X}\right)^2 - \frac{\hat{a}}{n}.$$

Ex:

policy	year 1	year 2	year 3	average	variance
1	3	5	7	5	4
2	6	12	9	9	9
				7	6.5

$= (3+5+7)/3$

$= (6+12+9)/3$ $\quad \cancel{(3-5)}^2$

$$= \frac{(3-5)^2 + (5-5)^2 + (7-5)^2}{2}$$

$$\hat{\mu} = 7 \; ; \; \hat{a} = 6.5 \; , \; \hat{v} = 8 - \frac{6.5}{3} = 5.83$$

$$k = \frac{\hat{a}}{\hat{v}} = \frac{6.5}{5.83} = 1.11$$

$$Z = \frac{n}{n+k} = \frac{3}{3+1.11} = \frac{3}{4.11} = 0.73$$

policy(1) $\leftarrow$ $P_c = 0.73 * 5 + 0.27 * 7 = 5.54$

policy2 $\leftarrow$ $P_c = 0.73 * 9 + 0.27 * 7 = 8.46.$