# Construction and Evaluation of Actuarial Models

**Rajapaksha Premarathna**

# Table of Contents

## Modeling



## Model Estimation

## Credibility

# Modeling

### Some definitions and Notations

Ground up loss – the actual loss amount prior to modifications. The loss random variable is denoted $X$ . We generally assume $X \geq 0$ .

Cost per loss – the amount paid by insurer. This includes the zero payments. Denoted $Y_L$

Cost per payment – the amount paid by insurer which includes only the non-zero payments made by the insurer. It is also called the excess loss random variable or the left truncated and shifted variable. Denoted $Y_p$

Severity distribution – the distribution loss amount or the cost to the insurer

Frequency distribution – the distribution of the number of losses, or amount paid per unit time

Policy Limit – Maximum amount paid by insurance policy for a single loss, denoted $u$ . If there is a deductible the policy limit is $u-d$

Ordinary Deductible – for loss amounts below deductible, $d$ the insurer pays 0 and for loss amounts above d the insurer pays the difference of the loss amount and deductible.

Maximum Covered Loss – the amount $u$ for which no additional benefits are paid. Denoted $u$

### Case 1: Policy Limit $u$

Amount paid by insurer = $X \wedge u = \begin{cases} X & X \leq u \\ u & X > u \end{cases} = \min(X, u)$

Note: $X \wedge u$ is also referred to as Limited loss random variable

Therefore the expected value is:

$$E\left[X \wedge u\right] = \int_0^u x \cdot f\left(x\right) dx + u \cdot \left[1 - F_X\left(u\right)\right] \text{ (for continuous)}$$

$$E\left[X \wedge u\right] = \sum_{x_j \leq u} x_j \cdot p\left(x_j\right) + u \cdot \left[1 - F_X\left(u\right)\right] \text{ (for discrete)}$$

$$E\left[X \wedge u\right] = \int_0^u \left[1 - F_X\left(x\right)\right] dx = \int_0^u S_x(x) dx \text{ (for discrete or continuous)}$$

Also

$$E\left[\left(X \wedge u\right)^{k}\right] = \int_{0}^{u} x^{k} \cdot f\left(x\right)dx + u^{k} \cdot \left[1 - F_{X}\left(u\right)\right]$$ (for continuous)

And

$$Var\left[X \wedge u\right] = E\left[\left(X \wedge u\right)^{2}\right] - \left(E\left[X \wedge u\right]\right)^{2}$$

## Case 2: With an Ordinary deductible

Amount paid by insurer = cost per loss = left censored and shifted random variable

$$Y_L = (X-d)_+ = \begin{cases} 0 & X \le d \\ X-d & X > d \end{cases} = \max(X-d, 0) = X - \left(X \wedge d\right)$$

Therefore the expected values are:

$$E\left[Y_L\right] = E\left[(X-d)_+\right] = \int_{d}^{\infty} \left(x-d\right) \cdot f\left(x\right)dx$$ (for continuous)

$$E\left[Y_L\right] = E\left[(X-d)_+\right] = \sum_{x_j > u} (x_j - d) \cdot p\left(x_j\right)$$ (for discrete)

$$E\left[Y_L\right] = E\left[(X-d)_+\right] = E\left[X\right] - E\left[X \wedge d\right] = \int_{d}^{\infty} \left[1 - F_X\left(x\right)\right]dx = \int_{d}^{\infty} S_x(x)dx$$ (for discrete or continuous)

$$Var\left[\left(X-d\right)_+\right] = E\left[\left(X-d\right)_+^{\ 2}\right] - \left(E\left[\left(X-d\right)_+\right]\right)^2 = Var\left(Y_L\right) = E\left(Y_L^2\right) - \left(E\left[Y_L\right]\right)^2$$

Also note that

$$E\left[Y_L^2\right] = E\left[(X-d)_+^2\right] = E\left[X^2\right] - E\left[\left(X \wedge d\right)^2\right] - 2d\left[E\left[X\right] - E\left[X \wedge d\right]\right]$$

When considering the expected cost per payment $E\left[X - d \middle| X > d\right] = E\left[Y_p\right] = e(d)$

$$E\left[Y_P\right] = \frac{E\left[Y_L\right]}{P\left(X > d\right)} = \frac{E\left[(X-d)_+\right]}{1 - F_X\left(d\right)} = \frac{E\left[X\right] - E\left[X \wedge d\right]}{1 - F\left(d\right)}$$

$$E\left[Y_P\right] = \frac{E\left[(X-d)_+\right]}{1 - F_X(d)} = \frac{\int_{d}^{\infty} \left(x-d\right) \cdot f\left(x\right)dx}{\int_{d}^{\infty} f\left(x\right)dx}$$ (for continuous)

$$E[Y_P] = \frac{E[(X-d)_+]}{1-F(d)} = \frac{\sum_{x_j>u} (x_j - d)\cdot p(x_j)}{\sum_{x_j>u} p(x_j)} \text{ (for discrete)}$$

$$E[Y_P] = \frac{\int_d^\infty [1-F_X(x)]dx}{1-F(d)} = \frac{\int_d^\infty S_x(x)dx}{S(d)} \text{ (for discrete or continuous)}$$

The expected cost per payment is also referred to as the mean excess loss, or the mean residual loss or mean residual lifetime.

Here are some shortcuts that will be useful in the examination:

1. Given $X$ is a uniform distribution on $[0,\theta]$ and an ordinary deductible $d$ is applied. Then:

$E[Y_p] = \frac{\theta - d}{2}$ and $E[Y_P^2] = \frac{(\theta-d)^2}{3}$ and the variance is $Var[Y_p] = \frac{(\theta-d)^2}{12}$

2. Given $X$ is an exponential distribution with mean $\theta$ and an ordinary deductible $d$ is applied. Then:

$E\lfloor Y_p \rfloor = \theta$ and $E\lfloor Y_P^2 \rfloor = 2\theta^2$ and the variance is $Var\lfloor Y_p \rfloor = \theta^2$

3. Given $X$ is a Pareto distribution with parameters $\alpha$ and $\theta$ and an ordinary deductible $d$ is applied $(\alpha > 1)$. Then:

The pdf of $Y_P$ is also Pareto with parameters $\alpha$ and $\theta' = d + \theta$. Therefore, $E[Y_P] = \frac{d+\theta}{\alpha - 1}$ and

$$E[Y_P^2] = \frac{2(d+\theta)^2}{(\alpha-1)(\alpha-2)}$$

4. Given $X$ is a single parameter Pareto distribution with parameters $\alpha$ and $\theta$ and an ordinary deductible $d$ is applied $(\alpha > 1)$.
   If $d \le \theta$ then $E[Y_P] = E[Y_L] = E[X] - d$ and $Var[Y_P] = Var[Y_L] = Var[X]$
   If $d > \theta$ then $Y_P$ has two parameter Pareto distribution with parameters $\alpha$ and $\theta = d$

The variance of cost per loss with a deductible $d$ is

$$Var[X-d|X>d] = Var[Y_P] = E[Y_P^2] - (E[Y_p])^2 = E[(X-d)^2|X>d] - (E[X-d|X>d])^2$$

Note that

$$E\left[Y_P^2\right]=\frac{E\left\lfloor Y_L^2\right\rfloor}{P\left(X>d\right)}=\frac{E\left\lfloor (X-d)_+^2\right\rfloor}{1-F_X\left(d\right)}=E\left[\left(X-d\right)^2\middle|X>d\right]$$

**Case 3: Maximum Covered loss $u$ with a policy deductible $d<u$**

Therefore the Cost per Loss $Y_L$ is

$$Y_L=\begin{cases}0 & X\le d\\ X-d & d<X\le u\\ u-d & X>u\end{cases}=\left(X\wedge u\right)-\left(X\wedge d\right)$$

The expected cost per loss is:

$$E\left[Y_L\right]=E\left[X\wedge u\right]-E\left[X\wedge d\right]=\int_d^u\left(x-d\right)\cdot f\left(x\right)dx+\left(u-d\right)\cdot\left[1-F_X\left(u\right)\right]=\int_d^u\left[1-F_X\left(x\right)\right]dx$$

The second moment for cost per loss therefore is:

$$E\left[Y_L^2\right]=\left(E\left[\left(X\wedge u\right)^2\right]-E\left[\left(X\wedge d\right)^2\right]\right)-2d\left(E\left[X\wedge u\right]-E\left[X\wedge d\right]\right)=\int_d^u\left(x-d\right)^2\cdot f\left(x\right)dx+\left(u-d\right)^2\cdot\left[1-F_X\left(u\right)\right]$$

The expected cost per payment is:

$$E\left[Y_P\right]=\frac{E\left[Y_L\right]}{1-F\left(d\right)}=\frac{E\left[X\wedge u\right]-E\left[X\wedge d\right]}{1-F\left(d\right)}$$

The second moment of cost per payment is:

$$E\left[Y_P^2\right]=\frac{E\left\lfloor Y_L^2\right\rfloor}{1-F\left(d\right)}$$

**Case 4: Coinsurance factor $\alpha$ where $0<\alpha<1$ and/or inflation rate $r$**

With maximum covered loss $u$ and deductible $d$ the amount paid by the insurer is (no inflation):

$$Y_L=\begin{cases}0 & X\le d\\ \alpha\left(X-d\right) & d<X\le u\\ \alpha\left(u-d\right) & X>u\end{cases}$$

The expected cost per loss is:

$$E[Y_L] = \alpha(E[X \wedge u] - E[X \wedge d]) = \alpha\int_d^u [1 - F_X(x)]dx$$

The expected cost per payment is:

$$E[Y_P] = \frac{E[Y_L]}{1 - F_X(d)} = \frac{\alpha(E[X \wedge u] - E[X \wedge d])}{1 - F_X(d)}$$

With inflation the expected cost per loss is:

$$E[Y_L] = \alpha(1+r)\left(E\left\lfloor X \wedge \frac{u}{1+r}\right\rfloor - E\left\lfloor X \wedge \frac{d}{1+r}\right\rfloor\right)$$

And the expected cost per payment is:

$$E[Y_P] = \frac{E[Y_L]}{1 - F_X\left(\frac{d}{1+r}\right)} = \frac{\alpha(1+r)\left(E\left\lfloor X \wedge \frac{u}{1+r}\right\rfloor - E\left\lfloor X \wedge \frac{d}{1+r}\right\rfloor\right)}{1 - F_X\left(\frac{d}{1+r}\right)}$$

**Other Concepts**

The Loss Elimination ratio is $\frac{E[X \wedge d]}{E[X]}$

**Bonus Payments**

If there is a bonus for loss amounts less than a specific limit $u$

$$\text{Bonus} = \begin{cases} u - X & 0 < X \le u \\ 0 & X > u \end{cases}$$

Therefore the expected bonus payment is $E[Bonus] = u - E[X \wedge u]$ and if the bonus is equal to a fraction $c$ of the amount by which the loss is less than $u$ then $E[Bonus] = c\lfloor u - E[X \wedge u]\rfloor$

## Franchise Deductible

Franchise deductible is when the insurer pays the full amount a deductible denoted $d$ . Therefore

$$\text{Amount paid by insurer} = \begin{cases} 0 & X \le d \\ X & X > d \end{cases}$$

Therefore the Expected cost per loss is $= \int_d^{\infty} x \cdot f_X(x) dx = E\left[(X-d)_+\right] + d\left[1 - F_X(d)\right]$ and the

Expected cost per payment =

$$E[Y_P] = \frac{E[Y_L]}{P(X>d)} = \frac{E\lfloor (X-d)_+ \rfloor + d\lfloor 1-F_X(d) \rfloor}{1-F_X(d)} = \frac{E\lfloor (X-d)_+ \rfloor}{1-F(d)} + d$$

## Compound Distributions

***Terminology***

$N$ Is the number of claims or the claim count random variable. The distribution is called the claim count distribution or frequency distribution.

$X$ Is the single or individual loss random variable whose distribution is known as the severity distribution

$S = X_1 + X_2 + \ldots + X_N$ and is the aggregate loss per period and has a compound distribution

$N, X_1, X_2, \ldots, X_N$ are mutually independent random variables

$E[S] = E[N] \cdot E[X]$

$Var\left[E\left[S|N\right]\right] = Var[N] \cdot \left(E[X]\right)^2$ and $E\left[Var\left[S|N\right]\right] = E[N] \cdot Var[X]$

If it is a compound Poisson distribution $S$ and the frequency distribution is Poisson with mean $\lambda$ then $E[S] = \lambda \cdot E[X]$ and $Var[S] = \lambda \cdot E\lfloor X^2 \rfloor$

***The distribution of $S$ if $N$ is in the $(a,b,1)$ class***

We know that $P(N=k) = p_k$ and $P(S=j) = f_S(j)$ . Also $P(X=x) = f_X(x)$

$$P(S=x) = f_S(x) = \frac{\left[p_1 - (a+b)\,p_0\right] \cdot f_X(x) + \sum_{j=1}^{x} \left(a + \frac{b}{x} \cdot j\right) \cdot f_X(j) \cdot f_S(x-j)}{1 - a \cdot f_X(0)}$$

***The distribution of $S$ if $N$ is in the $(a,b,0)$ class***

$$P(S=x)=f_S(x)=\frac{\sum_{j=1}^{x}\left(a+\frac{b}{x}\cdot j\right)\cdot f_X(j)\cdot f_S(x-j)}{1-a\cdot f_X(0)}$$

***The distribution of $S$ if $N$ is Poisson***

$$f_S(x)=\frac{\lambda}{x}\sum_{j=1}^{x} j\cdot f_X(j)\cdot f_S(x-j)$$

## Stop Loss Insurance

If a deductible is applied to aggregate losses the insurance payment is the aggregate loss in excess of the deductible. The stop loss insurance payment is

$$Max\{S-d,0\}=(S-d)_+=S-(S\wedge d)=\begin{cases}0 & S\le d\\ S-d & S>d\end{cases}$$

The expected value stop loss insurance payment is called the net stop loss premium where $E\lfloor (S-d)_+\rfloor=E[S]-E\lfloor (S\wedge d)\rfloor$

Also

$E\left[(S-d)_+\right]=E\left[(S-a)_+\right]-(d-a)\left[1-F_S(a)\right]=\frac{b-d}{b-a}\cdot E\left[(S-a)_+\right]+\frac{b-a}{b-a}\cdot E\left[(S-b)_+\right]$ where $a\le d\le b$ and $E\lfloor\lfloor S-(d+1)_+\rfloor\rfloor=E\lfloor (S-d)_+\rfloor-\lfloor 1-F_S(d)\rfloor$ for $d\ge 0$

# Model Estimation

## Review of Estimators

$E\left[\bar{X}\right] = \mu_X = E\left[X\right]\left(\bar{X}\text{ is an unbiased estimator of }\mu_X\text{ and }\bar{X}\text{ has the same mean as }X\right)$

$Var\left[\bar{X}\right] = \dfrac{\sigma_X^2}{N}\left(\text{this is the variance of the sample mean}\right)$

For a parameter $\beta$ the estimator is unbiased if $E\left\lfloor\hat{\beta}\right\rfloor = \beta$

The Bias of the parameter estimator is $Bias\left\lfloor\hat{\beta}\right\rfloor = E\left\lfloor\hat{\beta}\right\rfloor - \beta$

The Mean Square error or MSE is $E\left[\left(\hat{\beta}-\beta\right)^2\right] = \left(Bias\left[\hat{\beta}\right]\right)^2 + Var\left[\hat{\beta}\right]$

## Non-Parametric Empirical Point Estimation

The random variable $X$ can be a loss random variable or a failure time random variable. It can be discrete or continuous. A failure time random variable describes the time until a particular even happens.

Sample information for estimating the random variable $X$ is available in the in one of the following ways:

1. A random sample of independent n individual observations
2. Grouped data: the range of the random variable is divided to a series of intervals, $(-\infty, c_0)(c_0, c_1), \ldots\ldots, (c_{r-1}, c_r)(c_r, \infty)$ and the number of observations in an interval $(c_{j-1}, c_j)$ is $n_j$
3. Censored or truncated data

## Case 1: Empirical estimation from a random sample with complete individual data

If the exact values of n observations $x_1, x_2, \ldots., x_n$ (where $x_i$ is a loss amount given the data is a loss distribution or it is times of death or failure given it is survival distribution) the data is considered to be complete. A probability of $\dfrac{1}{n}$ is assigned for each $x_i$. If there are $k$ distinct numerical values such that these $k$ values or ordered from smallest to largest as $y_1 < y_2 < \ldots. < y_n$ with $s_j$ =number of observations equal to $y_j$ and $s_1 + s_2 + \ldots.. + s_n = 1$.

The empirical distribution probability function is $p_n(y_j) = \frac{\text{number of } x_i\text{'s that are equal to } y_j}{n} = \frac{s_j}{n}$

The empirical distribution function is $F_n(t) = \frac{\text{number of } x_i\text{'s} \le t}{n}$

The empirical survival function is $S_n(t) = 1 - F_n(t) = \frac{\text{number of } x_i\text{'s} > t}{n}$

The risk set at $y_j$ is denoted $r_j$, where $r_1 = n$. If there are $s_1$ deaths at time $y_1$ so there are $r_2 = n - s_1$ at risk at second death time $y_2$. If there are $s_2$ deaths at time $y_2$ so there are $r_2 = n - (s_1 + s_2)$ at risk at third death time.

The Nelson-Aelon estimate of the cumulative hazard function is

$$\widehat{H}(t) = \begin{cases} 0 & t < x_1 \\ \sum_{i=1}^{j-1} \frac{s_i}{r_i} & x_{j-1} \le t < x_j,\ j = 2,3,\dots,k \\ \sum_{i=1}^{k} \frac{s_i}{r_i} & x_k \le t \end{cases}$$

The Nelson-Aelon estimate of the survival function is $\widehat{S}(x) = e^{-\hat{H}(x)}$, and the Nelson-Aelon estimate of the distribution function is $\widehat{F}(x) = 1 - \widehat{S}(x) = 1 - e^{-\hat{H}(x)}$

In order to find the smoothed empirical estimate of the 100p-th percentile $\hat{\pi}_p$ use the following steps.

1. Order the sample values from smallest to largest.
2. Find an integer g such that $\frac{g}{n+1} \le p \le \frac{g+1}{n+1}$
3. $\hat{\pi}_p$ is found by linear interpolation $\hat{\pi}_p = \left[g + 1 - (n+1)p\right] x_{(g)} + \left[(n+1)p - g\right] x_{(g+1)}$

**Case 2: Empirical Estimation from Grouped Data**

The empirical estimate of the mean of $X$ is $\sum_{j=1}^{r} \left( \frac{n_j}{n} \cdot \frac{c_j + c_{j-1}}{2} \right)$. We assume the loss amounts are uniformly distributed. The empirical estimate of the k-th moment is: $\sum_{j=1}^{r} \left( \frac{n_j}{n} \cdot \frac{c_j^{k+1} - c_{j-1}^{k+1}}{(k+1)(c_j - c_{j-1})} \right)$

### Case 3: Estimation from Censored and truncated data

A truncated observation is data point that is not observed. Left truncation is truncation below (deductible). A censored observation is an observation that is observed to occur, but whose value is not known. Right censoring is censoring from above (policy limit).

Data description

If individual $i$ is a left truncated data point who has a value $d_i$ that satisfies $y_{j-1} \le d_i < y_j$, then we add that individual to the risk set $r_j$ for the next death point $y_j$ and individual $m$ is right censored data point who has a value $u_m$ that satisfies $y_{j-1} \le u_m < y_j$, then we remove that individual from the risk set $r_j$ for the next data point $y_j$. This is similar to the following:

$r_j = r_{j-1} - s_{j-1} + \left(\text{the number of individuals who have } y_{j-1} \le d_i < y_j\right) - \left(y_{j-1} \le u_m < y_j\right)$ where $s_j$ is the number of deaths at death point $y_j$

If truncated or censored observation time is the same as death time $y_j$ that individual is added or removed after the deaths at death point $y_j$ and it only affects the risk set $r_{j+1}$

Therefore:

$$r_j = \left(\text{number of } d\text{'s} < y_j\right) - \left(\text{number of } x\text{'s} < y_j\right) - \left(\text{number of } u\text{'s} < y_j\right) \text{ or}$$
$$r_j = \left(\text{number of } x\text{'s} \ge y_j\right) + \left(\text{number of u's} \ge y_j\right) - \left(\text{number of d's} \ge y_j\right)$$

### The Kaplan-Meier/Product Limit Estimator

$$S_n(t) = \begin{cases} 1 & 0 \le t < y_1 \\ \prod_{i=1}^{j-1}\left[1 - \frac{s_i}{r_i}\right] & y_{j-1} \le t < y_j,\ j = 2,3,\ldots,k \\ \prod_{i=1}^{k}\left[1 - \frac{s_i}{r_i}\right] \text{ or } 0 & t \ge y_k \end{cases}$$

If $z$ denoted the largest observation in the data set. Therefore when estimating $S(t)$ for $t > z$:

1. $S_n(t) = 0$
2. $S_n(t) = S_n(z)$

3. $S_n(t)=\left[S_n(z)\right]^{t/z}$ (geometric extension approximation)

## Kaplan Meier Approximation for Large data Sets

First choose a sequence of time points say $c_0 < c_1 < .... < c_k$. For an interval $(c_j, c_{j+1}]$, the number of uncensored observed deaths is denoted $x_j$; the number of right censored observations is denoted $u_j$ and $d_j$ denotes the number of left truncated observations.

The number at risk at time 0, $r_0 = d_0$. The number at risk for time interval $(c_j, c_{j+1}]$ is

$$r_j = \sum_{i=0}^{j} d_i - \sum_{i=0}^{j-1}(x_i + u_i)$$

The product limit estimate for the survival probability to the point $c_j$ is $\left(1-\frac{x_0}{r_0}\right)\left(1-\frac{x_1}{r_1}\right)\cdots\left(1-\frac{x_{j-1}}{r_{j-1}}\right)$

A variation on the Kaplan Meier/Product Limit large approach is defined by the following factors:

$$P_0 = 0$$

$P_j = \sum_{i=0}^{j-1}(d_i - u_i - x_i)$ and the number at risk at time $c_{j+1}$ is $r_j = P_j + \alpha d_j - \beta u_j$

## Variance of Survival Probability Estimates

If there is no censoring or truncation given individual data the empirical estimate of the survival function $S_n(x)=\frac{\text{number of deaths that occur after time x}}{n}=\frac{Y}{n}=\frac{n_x}{n}$ where $n_x$ =the number of survivors to time x

Also the estimator is an unbiased and consistent estimator of $S(x)$ and the variance is

$$Var\left[S_n(x)\right]=\frac{S(x)\left\lfloor 1-S(x)\right\rfloor}{n}$$

For grouped data for $n$ data points with intervals in the form $(c_0,c_1],(c_1,c_2],....,(c_{j-1},c_j],(c_j,\infty)$ the variance of $S_n(x)$ is

$$Var\left[S_n(x)\right]=\frac{\left(c_j-c_{j-1}\right)^2 Var\left[Y\right]+\left(x-c_{j-1}\right)^2 Var\left[m_j\right]+2\left(c_j-c_{j-1}\right)\left(x-c_{j-1}\right)Cov\left[Y,m_j\right]}{\left[n\left(c_j-c_{j-1}\right)\right]^2}$$

And

$$S_n(x) = \frac{c_j - x}{c_j - c_{j-1}} \cdot S_n(c_{j-1}) + \frac{x - c_{j-1}}{c_j - c_{j-1}} \cdot S_n(c_j)$$

Where

$x$ = value between interval $(c_{j-1}, c_j]$

$S_n(c_j) = \frac{n_j}{n}$

$Y = n - n_{j-1}$

$n_j$ = number of survivors at time $c_j$

$m_j$ = number of deaths in interval $(c_{j-1}, c_j]$

$Var[Y] = n \cdot S(c_{j-1}) \lfloor 1 - S(c_{j-1}) \rfloor$

$Var[m_j] = n \cdot \lfloor S(c_{j-1}) - S(c_j) \rfloor \cdot \lfloor 1 - S(c_{j-1}) + S(c_j) \rfloor$

$Cov[Y, m_j] = -n \cdot \lfloor 1 - S(c_{j-1}) \rfloor \cdot \lfloor S(c_{j-1}) - S(c_j) \rfloor$

The estimate for the density function in the interval $(c_{j-1} - c_j)$ is $f_n(x) = \frac{S_n(c_{j-1}) - S_n(c_j)}{c_j - c_{j-1}}$ and the variance of the estimator is $Var[f_n(x)] = \frac{\lfloor S_n(c_{j-1}) - S_n(c_j) \rfloor \cdot \lfloor 1 - S_n(c_{j-1}) + S_n(c_j) \rfloor}{n(c_j - c_{j-1})^2}$

The Greenwood's Approximation of the estimated variance of the product limit estimator is

$$V\hat{a}r[S_n(y_j)] = [S_n(y_j)]^2 \cdot \sum_{i=1}^{j} \frac{s_i}{r_i(r_i - s_i)}$$

The estimated variance of the Nelson Aalen estimate of the cumulative hazard function $H(y_j)$ is

$$V\hat{a}r[\hat{H}(y_j)] = \sum_{i=1}^{j} \frac{s_i}{r_i^2}$$

## Confidence Interval for Survival Probability Estimates

For an estimator $\hat{\theta}$ for a parameter $\theta$ the 95% linear confidence interval for $\theta$ is $\hat{\theta} \pm 1.96\sqrt{\text{var}\left(\hat{\theta}\right)}$

The lower limit for the 95% log transformed confidence interval for $S\left(t\right)$ is $S_n\left(t\right)^{1/U}$ and the upper limit is $S_n\left(t\right)^U$ where $U = \exp\left[\frac{1.96\sqrt{V\hat{a}r\left[S_n\left(t\right)\right]}}{S_n\left(t\right)\cdot\ln\left[S_n\left(t\right)\right]}\right]$

The lower limit for the 95% log transformed confidence interval for $H\left(t\right)$ is $\frac{\hat{H}\left(t\right)}{U}$ and the upper limit is $\hat{H}\left(t\right)\cdot U$ where $U = \exp\left[\frac{1.96\sqrt{V\hat{a}r\left[\hat{H}\left(t\right)\right]}}{\hat{H}\left(t\right)}\right]$

Note: 1.96 is found using the normal distribution table provided

## Method of Moments

For a distribution defined in terms of $r$ parameters $\left(\theta_1, \theta_2, ...., \theta_r\right)$ the method of moments estimator of the parameter values is found by solving the $r$ equations: theoretical j-th moment = empirical j-th moment, j=1, 2,…., r

If the estimator has only one parameter $\theta$ ,then solve for $\theta$ from the equation theoretical distribution first moment = empirical distribution first moment

If the distribution has two parameters $\theta_1$ and $\theta_2$ then we solve the following equations, $E\left[X|\theta\right] = \text{empirical estimate of } E\left[X\right]$ and $E\left[X^2|\theta\right] = \text{empirical estimate of } E\left[X^2\right]$ or theoretical distribution variance = empirical distribution variance.

## Method of Percentile Matching

Given a random sample or an interval grouped data sample and a distribution with $r$ parameters, choose $r$ percentile points $p_1, ...., p_r$ and set the distribution $p_i$'yth percentile equal to the empirical estimate for the $p_i$'th percentile. The $r$ parameter values are found by solving the system of equations.

### Maximum Likelihood Estimation (MLE) Definition

Maximum Likelihood Estimation is used to estimate the parameters in a parametric distribution. We are trying to find the distribution parameters that would maximize the density or the probability of the data set occurring. First we create the likelihood function $L(\theta)$ where $\theta$ is the parameter being estimated.

For individual data $L(\theta)=\prod_{j=1}^{n} f(x_j;\theta)$ for a random sample $x_1, x_2, ....., x_n$ and for grouped data

$L(\theta)=\prod_{j=1}^{r}\left[F(c_j;\theta)-F(c_{j-1};\theta)\right]^{n_j}$ for $r$ intervals where interval $(c_j, c_{j-1}]$ has $n_j$ observations

### Maximum Likelihood Estimation for Complete Data (No truncation or Censoring)

Use the following steps to find the maximum likelihood estimation

1. Find $L(\theta)$
2. Find log likelihood $l(\theta)=\ln L(\theta)$
3. Set $\frac{d}{d\theta} l(\theta)=0$
4. Solve for $\theta$

### Likelihood function for Loss data with policy limit $u$ (right censored data)

The likelihood function is $L(\theta)=\left[\prod_{j=1}^{n} f(x_j;\theta)\right]\cdot\left[1-F(u;\theta)\right]^m$ where $m$ is the number of limit payments equal to $u$ (losses greater than $u$ ) and there are $n$ payments below the limit.

### Likelihood function for Loss data with Policy Deductible $d$ (left truncated data)

Loss data can be available in 2 forms

1. Insurance payments >0 denoted $y_1, y_2, ....., y_k$
2. Actual loss amounts greater than the deductible, $x_1, x_2, ....., x_k$

This means that $x_i = y_i + d$ therefore

$$L(\theta)=\prod_{j=1}^{k}\frac{f(x_j;\theta)}{1-F(d;\theta)}=\prod_{j=1}^{k}\frac{f(y_j+d;\theta)}{1-F(d;\theta)}$$

## Likelihood function for Loss data with Policy Deductible $d$ and Maximum Covered Loss $u$

If there $n$ observed payments $y_1, y_2, ....., y_n$ that satisfy $0 < y_i < u - d$ and $n$ loss amounts $x_1, x_2, ....., x_n$ where $x_i = y_i + d$ the likelihood function is:

$$L(\theta) = \frac{\left\lfloor \prod_{j=1}^{n} f\left(x_j;\theta\right) \right\rfloor \cdot \left[1 - F\left(u;\theta\right)\right]^m}{\left[1 - F\left(d;\theta\right)\right]^{n+m}} = \frac{\left\lfloor \prod_{j=1}^{n} f\left(y_j + d;\theta\right) \right\rfloor \cdot \left[1 - F\left(u;\theta\right)\right]^m}{\left[1 - F\left(d;\theta\right)\right]^{n+m}}$$

Where $m$ is the amount of observed limit payments equal to $u - d$ therefore there will be $m$ corresponding losses $\geq u$

## Maximum Likelihood of Exponential Distribution with parameter $\theta$

For complete individual data without truncation or censoring the MLE estimator for parameter $\theta$ is the sample mean $\bar{x} = \frac{1}{n} \sum x_i$ for a random sample of observations $x_1, x_2, ....., x_n$

For an exponential distribution with a data set with $m$ limit payments and policy limit $u$ the MLE of the MLE for the mean of $X$ is $\hat{\theta} = \frac{\sum x_i + mu}{n} = \frac{\text{total of all payment amounts}}{\text{number of non censored payments}}$

For an exponential distribution with a policy deductible the MLE for the mean of ground up loss is (given data available was insurance payments $y_1, y_2, ....., y_k$ )

$$\hat{\theta} = \frac{\sum y_i}{n} = \frac{\text{total of all insurance payment amounts}}{\text{number of insurance payments}}$$

## Maximum Likelihood Estimation Shortcuts for Distributions in Exam C Table (given no Truncation or Censoring)

For a random sample $x_1, x_2, ....., x_n$ of the following distributions:

For an inverse exponential Distribution with parameter $\theta$ the MLE of $\theta$ is $\frac{n}{\sum \frac{1}{x_i}}$

For a Pareto distribution with parameters $\alpha, \theta$ where $\theta$ is given the MLE of $\alpha$ is

$$\frac{n}{\sum_{i=1}^{n} \ln\left(x_i + \theta\right) - n \ln\left(\theta\right)}$$

For a Weibull Distribution with parameters $\tau, \theta$ where $\tau$ is given the MLE for $\theta$ is $\hat{\theta} = \left( \frac{1}{n} \cdot \sum_{i=1}^{n} x_i^{\tau} \right)^{1/\tau}$

For a Inverse Pareto distribution with parameters $\alpha, \theta$ where $\theta$ is given the MLE of $\alpha$ is

$$\frac{n}{\sum_{i=1}^{n} \ln(x_i + \theta) - \sum \ln(x_i)}$$

For a Inverse Weibull Distribution with parameters $\tau, \theta$ where $\tau$ is given the MLE for $\theta$ is

$$\hat{\theta} = \left( \frac{n}{\sum_{i=1}^{n} \frac{1}{x_i^{\tau}}} \right)^{1/\tau}$$

For a Normal Distribution with mean $\mu$ and variance $\sigma^2$ the MLE of $\mu$ is $\hat{\mu} = \bar{x}$ ,the sample mean. For a Normal Distribution with mean $\mu$ and variance $\sigma^2$ the MLE of $\sigma^2$ is $\hat{\sigma}^2 = \frac{1}{n} \cdot \sum_{i=1}^{n} (x_i - \hat{\mu})^2$ ,the biased form of the sample variance.

For a Lognormal Distribution with parameters $\mu$ and $\sigma^2$ the MLE of $\mu$ is $\hat{\mu} = \frac{1}{n} \cdot \sum_{i=1}^{n} \ln(x_i)$,the sample mean. For a Normal Distribution with mean $\mu$ and variance $\sigma^2$ the MLE of $\sigma^2$ is $\hat{\sigma}^2 = \frac{1}{n} \cdot \sum_{i=1}^{n} (\ln x_i - \hat{\mu})^2$ ,the biased form of the sample variance.

For a Gamma distribution with parameters $\alpha, \theta$ where $\alpha$ is given the MLE of $\theta$ is $\hat{\theta} = \frac{1}{n\alpha} \cdot \sum x_i$

For a Inverse Gamma distribution with parameters $\alpha, \theta$ where $\alpha$ is given the MLE of $\theta$ is $\hat{\theta} = \frac{n\alpha}{\sum \frac{1}{x_i}}$

For a Poisson distribution with parameter $\lambda$ the MLE of $\lambda$ is $\hat{\lambda} = \frac{1}{n} \sum_{k=0}^{j} k n_k$ where the total number of observations is $n = n_0 + n_1 + .... + n_j$

For a Binomial Distribution with parameters $m, q$ if the sample variance is larger than the sample mean, then the binomial distribution is not a good fit for the data. If $m$ is known or given for a data set

$n_0, n_1, ... n_m$ he moment estimate and the MLE of $q$ are both

$$\hat{q} = \frac{1}{m} \frac{\sum_{k=0}^{m} k n_k}{\sum_{k=0}^{m} n_k} = \frac{\text{total number of heads}}{\text{total number of coin tosses}}$$

**Maximum Likelihood Estimation for Grouped Data**

The data given is grouped into 4 categories

Category 1: data value $x_i$ that has no truncation or censoring

Category 2: data value $u_i$ , no deductible but policy limit $u_i$

Category 3: data value $x_i$ before a deductible $d_i$ and no policy limit

Category 4: policy limit payment $u_i - d_i$ with deductible $d_i$ and maximum covered loss $u_i$

$C_1 = \text{the sum of } x_i\text{'s in Category 1}$

$C_2 = \text{the sum of u}_i\text{'s in Category 2}$

$C_3 = \text{the sum of } x_i - d_i\text{'s in Category 3}$

$C_4 = \text{the sum of } u_i - d_i\text{'s in Category 4}$

$n_i = \text{number of data points in Category } i\,(i = 1,2,3,4)$

The MLE for an exponential distribution with mean $\lambda$ is $\hat{\lambda} = \dfrac{C_1 + C_2 + C_3 + C_4}{n_1 + n_3}$

For a single parameter Pareto distribution with parameters $\alpha, \theta$ where $\theta$ is given:

Category 1: $z_i = \ln\left(\dfrac{x_i}{\theta}\right)$

Category 2: $v_i = \ln\left(\dfrac{u_i}{\theta}\right)$

Category 3: $w_i = \ln\left(\dfrac{x_i}{d_i}\right)$

Category 4: $y_i = \ln\left(\frac{u_i}{d_i}\right)$

The MLE of $\alpha$ is $\hat{\alpha} = \frac{n_1 + n_3}{C_1 + C_2 + C_3 + C_4}$ where the C and n factors are defined the same as earlier

For a Weibull Distribution with parameters $\tau, \theta$ where $\tau$ is given:

Category 1: $z_i = x_i^\tau$, for $x_i$ that is not censored or truncated

Category 2: $v_i = u_i^\tau$, for $u_i$ that is right censored(limit payment) and not truncated(no deductible)

The MLE for $\theta$ is $\hat{\theta} = \left(\frac{C_1 + C_2}{n_1}\right)^{1/\tau}$ and if the data is separated to four categories:

Category 1: $z_i = x_i^\tau$, for $x_i$ that is not censored or truncated

Category 2: $v_i = u_i^\tau$, for $u_i$ that is right censored(limit payment) and not truncated(no deductible)

Category 3: $w_i = x_i^\tau - d_i^\tau$

Category 4: $y_i = u_i^\tau - d_i^\tau$

The MLE for $\theta$ is $\hat{\theta} = \left(\frac{C_1 + C_2 + C_3 + C_4}{n_1 + n_3}\right)^{1/\tau}$

# Credibility

Given a random variable $X$ from a random sample $X_1, X_2, ..., X_n$ the goal of credibility theory is to estimate the mean of $X$

### Limited Fluctuation Credibility Theory

If the random sample being analyzed is $W$ and there is $n$ independent observations $W_1, W_2, ..., W_n$ available. Also the mean of $W$ is $\mu$ and the variance is $\sigma^2$ then full credibility standard is satisfied when $P\left[\left|\overline{W}-\mu\right| < k\mu\right] \geq P$ is satisfied where $k$ is some fraction of $\mu$ .

Range parameter *k*: usually k = 0.05

Probability Level *P*: usually P = 0.90

Full credibility standard is satisfied when $P\left[\left|\overline{W}-\mu\right| < k\mu\right] \geq P$ is satisfied

Once *P* and *k* are chosen we find a value *y* such that $P\left[-y \leq Z \leq y\right] = P$ where Z is the standard normal distribution. Therefore if $P = 0.90,$ then $y = 1.645$

Then chose $n_0 = \left(\frac{y}{k}\right)^2$

Therefore, for a random variable W, full credibility is given to $\overline{W}$ if the following conditions are satisfied

1. $n \geq n_0\left(\frac{Var(W)}{E(W)^2}\right) = n_0 \cdot$ (square of coefficient of variation) where n is the number of observations of W
2. The sum of all observed W values $\geq n_0 \frac{Var\left[W\right]}{E\left[W\right]}$

### Full Standard of Credibility for Compound distributions

Let compound distribution random variable be *S.* S has two components N (Frequency) and Y (Severity). Severity is a non-negative random variable that can be continuous or discrete. Usually S represents aggregate claims (per period) while N represents number of claims (per period or per policy holder) and Y represents size of claim.

We know that mean of variance of S is $E[S]=E[N]E[Y]$ and

$Var[S]=Var[N]\cdot(E[Y])^2+E[N]\cdot Var[Y]$

Therefore

1. Number of observations of S needed $n \geq n_0 \cdot \dfrac{Var[S]}{(E[S])^2}$
2. Sum of all observed S's $\geq n_0 \cdot \dfrac{Var[S]}{E[S]}$
3. Total number of observed claims $\geq n_0 \dfrac{Var[S]\cdot E[N]}{(E[S])^2} = n_0 \dfrac{Var[S]}{(E[Y])^2 \cdot E[N]}$

If S has a compound Poisson distribution i.e. N is Poisson with mean $\lambda$, N and Y are mutually independent and S has compound Poisson with mean $\lambda E[Y]$ therefore the standard of full credibility for S is

1. Number of observations of S needed $n \geq n_0 \cdot \dfrac{Var[S]}{(E[S])^2} = \dfrac{n_0}{\lambda}\left[1+\dfrac{Var[Y]}{(E[Y])^2}\right]$
2. Sum of all observed S's $\geq n_0 \cdot \dfrac{Var[S]}{E[S]} = n_0 \cdot \left[E[Y]+\dfrac{Var[Y]}{E[Y]}\right]$
3. Total number of observed claims $\geq n_0 \dfrac{Var[S]\cdot E[N]}{(E[S])^2} = n_0 \cdot \left[1+\dfrac{Var[Y]}{E[Y]}\right]$

**Full Credibility Standard for Poisson Random Variable N (Number of Claims)**

1. Number of observed values of N needed = number of periods needed $n \geq n_0 \left(\dfrac{Var(N)}{E(N)^2}\right) = \dfrac{n_0}{\lambda}$
2. Total number of claims needed $\geq n_0$

**Partial Credibility**

The credibility premium $P = Z\bar{W} + (1-Z)M$ where $\bar{W}$ is the sample mean and M is the manual premium. Z is called the credibility factor where

$$Z = \sqrt{\frac{\text{info available}}{\text{info needed for full credibility}}}$$

For example to satisfy condition 1 the partial credibility factor is

$$Z=\sqrt{\frac{number\quad of\quad observations\quad available}{number\quad of\quad observations\quad needed\quad for\quad full\quad credibility}}$$

$P[A]=P[A\cap B]+P[A\cap B']=P\left[A|B\right]\cdot P[B]+P\left[A|B'\right]\cdot P[B']$ and

$$P\left[B|A\right]=\frac{P[A\cap B]}{P[A]}=\frac{P\left[A|B\right]\cdot P[B]}{P\left[A|B\right]\cdot P[B]+P\left[A|B'\right]\cdot P[B']}$$

**Predictive Probability**

$$P\left[B|A\right]=P\left[B|C\right]\cdot P\left[C|A\right]+P\left[B|C'\right]\cdot P\left[C'|A\right]$$

$$E[Y]=P\left[Y|C\right]\cdot P[C]+P\left[Y|C'\right]\cdot P[C']$$

$$E\left[Y|B\right]=P\left[Y|C\right]\cdot P\left[C|B\right]+P\left[Y|C'\right]\cdot P\left[C'|B\right]$$

$E[Y]=\sum_{j=1}^{m}E\left[Y|C_j\right]\cdot P\left[C_j\right]$ and $E\left[Y|B\right]=\sum_{j=1}^{m}E\left[Y|C_j\right]\cdot P\left[C_j|B\right]$

The initial assumption for the distribution (with parameter $\Theta$ ) is called the prior distribution and the pdf/pf is denoted $\pi(\theta)$. The distribution can be continuous or discrete.

The model distribution X is a conditional distribution (given $\Theta=\theta$ ) with pdf/pf $f_{X|\Theta}\left(x|\Theta=\theta\right)$. For a data set of random observed values from distribution of X and a specific $\theta$, the model distribution is

$$f_{X|\Theta}\left(x_1,x_2,\ldots,x_n|\Theta=\theta\right)=\prod_{i=1}^{n}f\left(x_i|\theta\right)=f\left(x_1|\theta\right)\cdot f\left(x_2|\theta\right)\cdots f\left(x_n|\theta\right)$$

The Joint distribution of X and $\Theta$ has pf/pdf $f_{X,\Theta}(x,\theta)=f\left(x|\theta\right)\cdot\pi(\theta)$ and for a data set

$$f_{X,\Theta}(x_1,x_2,\ldots,x_n,\theta)=f\left(x_1|\theta\right)\cdot f\left(x_2|\theta\right)\cdots f\left(x_n|\theta\right)\cdot\pi(\theta)$$

The marginal distribution of X is $f_X(x)=\int f\left(x|\theta\right)\cdot\pi(\theta)$ and for a data set

$f_X(x_1,x_2,\ldots,x_n)=\sum f\left(x_1|\theta\right)\cdot f\left(x_2|\theta\right)\cdots f\left(x_n|\theta\right)\cdot\pi(\theta)$ (for continuous $\Theta$ ) and

$f_X(x_1,x_2,\ldots,x_n)=\int f\left(x_1|\theta\right)\cdot f\left(x_2|\theta\right)\cdots f\left(x_n|\theta\right)\cdot\pi(\theta)d\theta$ (for discrete $\Theta$ )

The posterior distribution of $\Theta$ given $X = x$ has pdf/pf $\pi_{\Theta|X}\left(\theta|x\right) = \dfrac{f_{X,\Theta}\left(x,\theta\right)}{f_X\left(x\right)}$

Given data $x_1, x_2, \ldots, x_n$ , the predictive distribution of $X_{n+1}$ has pdf/pf

$f_{X_{n+1}|X}\left(x_{n+1}|x_1, x_2, \ldots, x_n\right) = \int f_{X_{n+1}|\Theta}\left(x_{n+1}|\theta\right) \cdot \pi_{\Theta|X}\left(\theta|x_1, x_2, \ldots, x_n\right) d\theta$ for continuous and

$f_{X_{n+1}|X}\left(x_{n+1}|x_1, x_2, \ldots, x_n\right) = \sum f_{X_{n+1}|\Theta}\left(x_{n+1}|\theta\right) \cdot \pi_{\Theta|X}\left(\theta|x_1, x_2, \ldots, x_n\right)$ for discrete

**Bayesian Credibility Shortcuts**

1. If model distribution is exponential with mean $\lambda$ and the prior distribution in inverse gamma with parameters $\alpha, \theta$ then:

   When a single data value is given the mean of marginal distribution of X is $\dfrac{\theta}{\alpha - 1}$ and the posterior distribution is inverse gamma with parameters $\alpha' = \alpha + 1$ and $\theta' = \theta + x$ and the predictive mean is $\dfrac{\theta + x}{\alpha}$

   When there are n data values the posterior distribution is inverse gamma with parameters $\alpha' = \alpha + n$ and $\theta' = \theta + \sum x_i$ and the predictive mean is $\dfrac{\theta + \sum x_i}{\alpha + n}$

2. If model distribution is Poisson with mean $\lambda$ and the prior distribution in gamma with parameters $\alpha, \theta$ then:

   When a single data value is given the mean of marginal distribution of X is negative binomial with $r = \alpha$ and $\beta = \theta$ and the posterior distribution is gamma with parameters $\alpha' = \alpha + x$ and $\theta' = \dfrac{\theta}{\theta + 1}$

   When there are n data values the posterior distribution is gamma with parameters $\alpha' = \alpha + \sum x_i$ and $\theta' = \dfrac{\theta}{n\theta + 1}$

   In both cases the predictive distribution is negative binomial with $r' = \alpha + \sum x_i$ and $\beta' = \dfrac{\theta}{n\theta + 1}$ and the predictive mean is the same as the mean of the posterior distribution

   If $\alpha = 1$ in the prior distribution the prior distribution becomes exponential with the marginal distribution X becomes geometric.

3. If the model distribution is binomial with parameters m, q and the prior distribution is beta with parameters a,b,1 then:

When a single data value is given the posterior distribution is beta with parameters a+x and b+m-x

When n data values are given the posterior distribution is beta with parameters $a+\sum x_i$ and $b+nm-\sum x_i$ and the predictive mean will be $m\times$(posterior mean)

4. If model distribution is inverse exponential with parameter $\lambda$ and the prior distribution in gamma with parameters $\alpha,\theta$ then:

   When a single data value is given the marginal distribution of X is inverse Pareto with $r=\alpha$ and the same $\theta$ and the posterior distribution is gamma with parameters $\alpha'=\alpha+1$ and $\frac{1}{\theta'}=\frac{1}{\theta}+\frac{1}{x}$ and the predictive mean is $\frac{\theta+x}{\alpha}$

   When there are n data values the posterior distribution is gamma with parameters $\alpha'=\alpha+n$ and $\frac{1}{\theta'}=\frac{1}{\theta}+\sum\frac{1}{x_i}$

5. If the model distribution is Normal with mean $\lambda$ and variance $\sigma^2$ and the prior distribution is Normal with mean $\mu$ and variance $\alpha$ then:

   For a single data value of x, the posterior distribution is normal with mean $\left(\frac{x}{\sigma^2}+\frac{\mu}{\alpha}\right)\Big/\left(\frac{1}{\sigma^2}+\frac{1}{\alpha}\right)$ and variance $\frac{1}{\left(\frac{1}{\sigma^2}+\frac{1}{\alpha}\right)}$

   For n data values the posterior distribution for $\lambda$ is Normal with mean $\left(\frac{\sum x_i}{\sigma^2}+\frac{\mu}{\alpha}\right)\Big/\left(\frac{1}{\sigma^2}+\frac{1}{\alpha}\right)$

   and variance $\frac{1}{\left(\frac{n}{\sigma^2}+\frac{1}{\alpha}\right)}$. Also the predictive mean is the same as the posterior mean.

6. If the model distribution is Uniform with on the interval $[0,\lambda]$ and the prior distribution is single parameter Pareto with parameters $\alpha,\theta$ then:

   If there are n observations $x_1,x_2,\ldots,x_n$ and M = $\max(x_1,x_2,\ldots,x_n,\theta)$ then the posterior distribution is single parameter Pareto with $\alpha'=\alpha+n$ and $\theta'=M$ the Bayesian premium is

   $\frac{(\alpha+n)M}{2(\alpha+n-1)}$

## Buhlmann Credibility

The initial structure for Buhmann credibility is the same as Bayesian credibility model. Therefore the model distribution $X$ is a conditional distribution (given $\Theta = \theta$) with pdf/pf $f_{X|\Theta}\left(x|\Theta=\theta\right)$The initial assumption for the distribution (with parameter $\Theta$ )is called the prior distribution and the pdf/pf is denoted $\pi\left(\theta\right)$. The distribution can be continuous or discrete.

Under Buhlmann credibility the conditional distributions of $X_i$'s given $\Theta=\theta$ is considered to be i.i.d (independent and identically distributed). Therefore:

$E\left[X_i|\Theta=\theta\right]=\mu\left(\theta\right)$ is the hypothetical mean

$Var\left[X_i|\Theta=\theta\right]=v\left(\theta\right)$ is the process variance

$E\left(X\right)=E\left[E\left[X|\Theta\right]\right]=E\left[\mu\left(\Theta\right)\right]=\mu$ is the pure premium or collective premium

$Var\left[E\left[X_i|\Theta\right]\right]=Var\left[\mu\left(\Theta\right)\right]=a$ is the variance of the hypothetical mean VHM

$E\left[Var\left[X_i|\Theta\right]\right]=E\left[v\left(\Theta\right)\right]=v$ is the expected process variance or EPV

Also

$Var\left[X_i\right]=v+a$

From this we calculate the Buhlmann Credibility Premium to be $Z\bar{X}+(1-Z)\mu$ where $Z=\dfrac{n}{n+k}$ where $k=\dfrac{v}{a}$. $Z$ is called the Buhlmann Credibility factor. If $a=0$ then $Z=0$.

## The Buhlmann Straub model

The difference between the original Buhlmann model and the Buhlmann Straub model is that the conditional variances of $X_i$ given $\Theta=\theta$ might not be the same. Therefore for a given measuring exposure $m_i$ where $m=m_1+m_2+....+m_n$ the process variance $Var\left[X_i|\Theta=\theta\right]=\dfrac{v\left(\theta\right)}{m_i}$

$E\left(X_i\right)=E\left\lfloor\mu\left(\Theta\right)\right\rfloor=\mu$ is the pure premium or collective premium

$Var\left\lfloor\mu\left(\Theta\right)\right\rfloor=a$ is the variance of the hypothetical mean VHM

$E\left[v(\Theta)\right] = v$ is the expected process variance or EPV

Also $Var\left[X_i\right] = a + \frac{v}{m_i}$

**Empirical Bayes Credibility Methods**

Our objective is still to apply the Buhlmann or Buhlmann-Straub models to determine the credibility premium based on observed claim data only using the following information:

1. Insurance Portfolio has $r$ policy holders where $i = 1, 2, 3...., r$
2. For policy holder $i$ data on $n_i$ exposure periods is available where $j = 1, 2, 3....n_i$
3. For policyholder $i$ and exposure period $j$ there are $m_{ij}$ exposure units with an average observed claim of $X_{ij}$ per exposure unit
4. The total claim observed for policyholder $i$ in exposure period $j$ is $m_{ij}X_{ij}$ and the total claim observed for policy holder $i$ in all $n_i$ exposure periods is $\sum_{j=1}^{n_i} m_{ij}X_{ij}$
5. The total number of exposure units for policyholder $i$ is $m_i = \sum_{j=1}^{n_i} m_{ij}$
6. The average observed claims per exposure unit for policyholder $i$ is $\bar{X}_i = \frac{1}{m_i}\sum_{j=1}^{n_i} m_{ij}X_{ij}$
7. The total number of exposure units for all policyholders is $m = \sum_{i=1}^{r} m_i$
8. The average claim per exposure period for all policyholders is
$$\bar{X} = \frac{1}{m}\sum_{i=1}^{r} m_i\bar{X}_i = \frac{\text{total observed claims for all policyholders in all periods}}{\text{total number of exposure periods for all policyholders}}$$
9. Policyholder $i$ has risk parameter variable $\Theta_i$ where each $\Theta_i$ is iid.
10. $E\left[X_{ij} \mid \Theta_i = \theta_i\right] = \mu(\theta_i)$ and $Var\left[X_{ij} \mid \Theta_i = \theta_i\right] = \frac{v(\theta_i)}{m_{ij}}$
11. $E\left[\mu(\Theta_i)\right] = \mu$, $Var\left[\mu(\Theta_i)\right] = a$ and $E\left[v(\Theta_i)\right] = v$
12. The credibility premium for the next exposure period for policyholder $i$ is $Z_i\bar{X}_i + (1 - Z_i)\mu$ where $Z_i = \frac{m_i}{m_i + \frac{v}{a}}$

**Empirical Bayes Estimation for Buhlmann model (Equal Sample Size)**

Since $n_1 = n_2 = .... = n_r = n$ and $m_{ij} = 1$ under the Buhlmann model use the following information to find the premium:

1. $\bar{X}_i = \frac{1}{n}\sum_{j=1}^{n} X_{ij}$ and $\bar{X} = \frac{1}{r}\sum_{j=1}^{r} \bar{X}_i = \frac{1}{r \cdot n}\sum_{i=1}^{r}\sum_{j=1}^{n} X_{ij}$
2. The estimated prior mean $\hat{\mu} = \bar{X}$
3. The estimated process variance is $\hat{v} = \frac{1}{r(n-1)}\sum_{i=1}^{r}\sum_{j=1}^{n}\left(X_{ij} - \bar{X}_i\right)^2 = \frac{1}{r}\sum_{i=1}^{r}\hat{v}_i$ where

   $\hat{v}_i = \frac{1}{n-1}\sum_{j=1}^{n}\left(X_{ij} - \bar{X}_i\right)^2$
4. The estimated variance of the hypothetical mean is $\hat{a} = \frac{1}{r-1}\sum_{i=1}^{r}\left(\bar{X}_i - \bar{X}\right)^2 - \frac{\hat{v}}{n}$
5. $\widehat{Z}_i = \frac{m_i}{m_i + \frac{\hat{v}}{\hat{a}}} = \frac{n}{n + \frac{\hat{v}}{\hat{a}}}$ and if $\hat{a} \le 0$ then $\hat{Z} = 0$

**Empirical Bayes Estimation for the Buhlmann Straub Model (Unequal Sample Size)**

Use the following information to find premium:

1. $\hat{\mu} = \bar{X}$
2. $\hat{v} = \frac{1}{\sum_{i=1}^{r}(n_i - 1)}\sum_{i=1}^{r}\sum_{j=1}^{n_i} m_{ij}\left(X_{ij} - \bar{X}_i\right)^2$
3. $\hat{a} = \frac{1}{m - \frac{1}{m}\sum_{i=1}^{r} m_i^2}\left[\sum_{i=1}^{r} m_i\left(\bar{X}_i - \bar{X}\right)^2 - \hat{v}(r-1)\right]$
4. $\widehat{Z}_i = \frac{m_i}{m_i + \frac{\hat{v}}{\hat{a}}} = \frac{n}{n + \frac{\hat{v}}{\hat{a}}}$

**Semi parametric Empirical Bayesian Credibility**

Consider a case where $X$ represents the number of claims for a period of time. Use the following information to find the premium:

1. $E\left[X\,|\Theta\right]=\mu(\theta)=\Theta$
2. $Var\left[X\,|\Theta\right]=v(\theta)=\Theta$
3. $\mu=v$
4. $Var\left[X\right]=v+a$
5. From this we calculate the Buhlmann Credibility Premium to be $Z\bar{X}+(1-Z)\mu$ where $Z=\frac{n}{n+k}$ where $k=\frac{v}{a}$. $Z$ is called the Buhlmann Credibility factor