EDUCATION COMMITTEE

OF THE

SOCIETY OF ACTUARIES

SHORT-TERM ACTUARIAL MATHEMATICS STUDY NOTE

TOPICS IN CREDIBILITY THEORY

by

Curtis Gary Dean, FCAS, MAAA

STAM-23-18 **Printed in U.S.A.**

Preface

This study note was written to supplement the “Credibility” chapter of *Foundations of Casualty Actuarial Science* as a reading for the fourth CAS/SOA examination. It presents important topics not covered in the *Foundations* chapter including the Bühlmann-Straub Model and nonparametric and semiparametric estimation of credibility formula parameters.

The author would like to thank Clive Keatinge, a member of both the CAS Examination and Syllabus Committees, for his suggestion to write the study note. Clive provided oversight during the writing and editing process, and made helpful suggestions about the content and clarity of exposition. He also chaired the committee that reviewed the study note. I would also like to thank the reviewers on the committee for their many valuable contributions including Joseph Boor, Russell Greig, Nasser Hadidi, Leigh Halliwell, Stuart Klugman, Walter Lowrie, Marjorie Rosenberg, and Gary Venter. Gary Venter is responsible for much of the material included in the section “An Intuitive Model for Credibility.”

Ball State University students also provided valuable comments and should be recognized: Heather Adams, Chuket Ounjitti, Doug Pirtle, Lori Thompson, and Melany Tower.

Table of Contents

1. Credibility Models

This study note supplements the "Credibility" chapter of *Foundations of Casualty Actuarial Science* as a reading for the fourth CAS/SOA examination. Several important topics not covered in the *Foundations* text are presented here, including the Bühlmann-Straub credibility model and estimation of credibility formula parameters. It is assumed that the student already has some familiarity with the material covered in the "Credibility" chapter before reading this study note.

The credibility models that will be discussed are often referred to as greatest accuracy credibility or least squares credibility. As will be explained later, these methods attempt to produce linear estimates that will minimize the expected value of the square of the difference between the estimate and the quantity being estimated.

Bühlmann credibility will be reviewed paying particular attention to the simplifying assumptions that distinguish it from the more general Bühlmann-Straub model that follows. The second half of the study note covers estimation of credibility formula parameters when underlying distributions are unknown.

Before beginning a more rigorous study, an intuitive derivation of the useful Bühlmann credibility model will be presented.

<u>An Intuitive Model for Credibility</u>

The actuary uses observations of events that happened in the past to forecast future events or costs. For example, data that was collected over several years about the average cost to insure a selected risk, sometimes referred to as a policyholder or insured, may be used to estimate the expected cost to insure the same risk in future years. Because insured losses arise from random occurrences, however, the actual costs of paying insurance losses in past years may be a poor estimator of future costs.

Consider a risk that is a member of a particular class of risks. Classes are groupings of risks with similar risk characteristics, and though similar, each risk is still unique and not quite the same as other risks in the class. In class rating, the insurance premium charged to each risk in a class is derived from a rate common to the class. Class rating is often supplemented with experience rating so that the insurance premium for an individual risk is based on both the class rate and actual past loss experience for the risk. The important question in this case is: How much should the class rate be modified by experience rating? That is, how much credibility should be given to the actual experience of the individual risk?

Intuition says that two factors appear important in finding the right balance between class rating and individual risk experience rating:

(1) **How homogeneous are the classes?** If all of the risks in a class are identical and have the same expected value for losses, then why bother with individual

experience rating? Just use the class rate. On the other hand, if there is significant variation in the expected outcomes for risks in the class, then relatively more weight should be given to individual risk loss experience.

Each risk in the class has its own individual risk mean called its hypothetical mean. The Variance of the Hypothetical Means (*VHM*) across risks in the class is a statistical measure for the homogeneity or vice versa, heterogeneity, within the class. A smaller *VHM* indicates more class homogeneity and, consequently, argues for more weight going to the class rate. A larger *VHM* indicates more class heterogeneity and, consequently, argues for less weight going to the class rate.

(2) **How much variation is there in an individual risk's loss experience?** If there is a large amount of variation expected in the actual loss experience for an individual risk, then the actual experience observed may be far from its expected value and not very useful for estimating the expected value. In this case, less weight, i.e., less credibility, should be assigned to individual experience. The process variance, which is the variance of the risk's random experience about its expected value, is a measure of the variability in an individual risk's loss experience. The Expected Value of the Process Variance (*EPV*) is the average value of the process variance over the entire class of risks.

Let $\overline{X}_i$ represent the sample mean of *n* observations for a randomly selected risk *i*. Because there are *n* observations, the variance in the sample mean $\overline{X}_i$ is the variance in one observation for the risk divided by *n*. Given risk *i*, this variance is PV_i / n where PV_i is the process variance of one observation. Because risk *i* was selected at random from the class of risks, an estimator for its variance is $E[\ PV_i / n\ ] = E[\ PV_i\ ] / n = EPV / n$. This is the Expected Value of the Process Variance for risks in the class divided by the number of observations made about the selected risk.[1] It measures the variability expected in an individual risk's loss experience.

Letting μ represent the overall class mean, a risk selected at random from the class will have an expected value equal to the class mean μ. The variance of the individual risk means about μ is the *VHM*, the Variance of the Hypothetical Means.

There are two estimators for the expected value of the i^{th} risk: (1) the risk's sample mean $\overline{X}_i$, and (2) the class mean μ. How should these two estimators be weighted together? A linear estimate with the weights summing to 1.00 would be

$$Estimate = w\overline{X}_i + (1-w)\mu \ .$$

An optimal method for weighting two estimators is to choose weights proportional to the reciprocals of their respective variances. This results in giving more weight to the

[1] The expectation is taken over all risks in the class.

estimator with smaller variance and less weight to the estimator with larger variance. In many situations this will result in a minimum variance estimator. (Please see the first problem in the exercises at the end of the study note.)

The resulting weights are

$$w = \frac{\dfrac{1}{EPV/n}}{\dfrac{1}{EPV/n} + \dfrac{1}{VHM}} \qquad \text{and} \qquad (1-w) = \frac{\dfrac{1}{VHM}}{\dfrac{1}{EPV/n} + \dfrac{1}{VHM}} \quad .$$

Note that a denominator was chosen so that the weights add to one. A little algebra produces

$$w = \frac{n}{n + \dfrac{EPV}{VHM}} \qquad \text{and} \qquad (1-w) = 1 - \frac{n}{n + \dfrac{EPV}{VHM}} \quad .$$

Setting $K = EPV / VHM$, the weight assigned to the risk's observed mean is

$$w = \frac{n}{n+K} \quad .$$

This is the familiar Bühlmann credibility formula with credibility $Z = n / (n + K)$.[2]

In this section, a risk selected from a rating class was used to illustrate the concept of credibility. In general, an individual risk or a group of risks comes from a larger population and the goal is to find the right balance between using the data for the smaller group and the larger population. Many other examples are possible.

Example An actuary calculated indicated rate changes by territory for automobile insurance. The rate change indication for the i^{th} territory was R_i. Combined data for the entire state indicated that a rate change of +2.0% was required. From these values, credibility weighted rate change indications were calculated:

$$\begin{array}{c} \textit{Credibility weighted rate change} \\ \textit{indication for territory } i \end{array} \quad = \quad Z_i \text{ x } R_i \; + \; (1 - Z_i) \text{ x } (+2.0\,\%) \ .$$

The credibility weights Z_i were calculated from the formula $Z_i = n_i / (n_i + K)$ where n_i was the number of insured vehicles in the territory during the three-year data collection period.

[2] A rigorous derivation of the Bühlmann credibility formula is provided in Appendix A.

Preliminaries and Notation

The actuary uses observations for a risk or group of risks to estimate future outcomes for that same risk or group. In this study note, although the term “a risk” is often used, the same comments can generally be applied to a group of risks where the group is a collection of risks with some common characteristics. The actual observation during time t for that particular risk or group will be denoted by x_t, which will be the observation of corresponding random variable X_t, where t is an integer. For example, X_t may represent the following:

- Number of claims in period t
- Loss ratio in year t
- Loss per exposure in year t
- Outcome of the t^{th} roll of a die.

An individual risk is a member of a larger population and the risk has an associated risk parameter θ that distinguishes the individual’s risk characteristics. It is assumed that the risk parameter is distributed randomly through the population and Θ will denote the random variable. The distribution of the random variable X_t depends upon the value of θ: $f_{X|\Theta}(x_t \mid \theta)$. For example, θ may be a parameter in the distribution function of X_t. In the case of a Poisson claims process, θ might be the expected number of claims. Although the examples in this study note will use θ’s that are scalars, one can also build models with $\boldsymbol{\theta}$ as a multidimensional vector with each component of the vector describing some aspect of the individual’s risk characteristics.

If X_t is a continuous random variable, the mean for X_t given $\Theta = \theta$, is the conditional expectation,

$$E_{X|\Theta}[X_t \mid \Theta = \theta] = \int x_t \, f_{X|\Theta}(x_t \mid \theta)dx_t = \mu(\theta) ,$$

where the integration is over the support of $f_{X|\Theta}(x_t \mid \theta)$. If X_t is a discrete random variable, then a summation should be used:

$$E_{X|\Theta}[X_t \mid \Theta = \theta] = \sum_{all\ x_t} x_t \, f_{X|\Theta}(x_t \mid \theta) .$$

The integral notation will be used in general cases, but the reader should be aware that a summation is called for with discrete random variables. It will be assumed that $\mu(\theta) = E_{X|\Theta}[X_t \mid \Theta = \theta]$ is constant through time for the models considered in this study note.[3]

The risk parameter represented by the random variable Θ has its own probability density function (p.d.f.): $f_{\Theta}(\theta)$. The p.d.f. for Θ describes how the risk characteristics are

[3] This is a major assumption that is easily violated in practice. Risk characteristics can change for a variety of reasons: a young driver becomes a better driver with experience; a business may institute risk control procedures that reduce losses; traffic densities may increase in an area leading to increased probabilities of auto accidents; and inflation will increase the costs of loss payments.

distributed within the population. If two risks have the same parameter θ, then they are assumed to have the same risk characteristics including the same mean $\mu(\theta)$.

The unconditional expectation of X_t is

$$E[X_t] = \iint x_t f_{X,\Theta}(x_t,\theta)dx_t d\theta = \iint x_t f_{X|\Theta}(x_t \mid \theta) f_\Theta(\theta) dx_t d\theta$$ [4]

$$= \int [\int x_t f_{X|\Theta}(x_t \mid \theta)dx_t] f_\Theta(\theta)d\theta = E_\Theta[E_{X|\Theta}[X_t \mid \Theta]] = E_\Theta[\mu(\theta)] = \mu .$$

The conditional variance of X_t given $\Theta = \theta$ is

$$Var_{X|\Theta}[X_t \mid \Theta = \theta] = E_{X|\Theta}[(X_t - \mu(\theta))^2 \mid \Theta = \theta] = \iint (X_t - \mu(\theta))^2 f_{X|\Theta}(x_t \mid \theta)dx_t = \sigma^2(\theta) .$$

This variance is often called the process variance for the selected risk. The unconditional variance of X_t, also referred to as the total variance, is given by the Total Variance formula:

$$Var[X_t] = Var_\Theta[E_{X|\Theta}[X_t \mid \Theta]] + E_\Theta[Var_{X|\Theta}[X_t \mid \Theta]], \text{ or}$$

Total Variance	=	*Variance of the Hypothetical Means*	+	*Expected Value of the Process Variance*

A proof of this formula is shown in Appendix B. These concepts are best demonstrated with an example.

Example The number of claims X_t during the t^{th} period for a risk has a Poisson distribution with parameter θ: $P[X_t = x] = \dfrac{\theta^x e^{-\theta}}{x!}$. The risk was selected at random from a population for which Θ is uniformly distributed over the interval [0,1]. (This simple distribution for Θ was chosen to make the integration easy.) It will be assumed that θ is constant through time for each risk.

(1) Hypothetical mean for risk with parameter θ is $\mu(\theta) = E_{X|\Theta}[X_t \mid \Theta = \theta] = \theta$ because the mean of the Poisson random variable is the parameter θ.

(2) Process variance for risk with parameter θ is $\sigma^2(\theta) = Var_{X|\Theta}[X_t \mid \Theta = \theta] = \theta$ because the variance equals the parameter θ for the Poisson.

(3) Variance of the Hypothetical Means (*VHM*) is

[4] Note that a substitution for the joint density function $f_{X,\Theta}(x_t,\theta)$ was made using the relationship $f_{X,\Theta}(x_t,\theta) = f_{X|\Theta}(x_t \mid \theta) f_\Theta(\theta)$.

$$Var_{\Theta}[E_{X|\Theta}[X_t|\Theta]] = Var_{\Theta}[\Theta] = E_{\Theta}[\Theta^2] - (E_{\Theta}[\Theta])^2 = \int_0^1 \theta^2(1)d\theta - \left(\int_0^1 \theta(1)d\theta\right)^2 = 1/12 \ .$$

(4) Expected Value of the Process Variance (*EPV*) is

$$E_{\Theta}[Var_{X|\Theta}[X_t|\Theta]] = E_{\Theta}[\Theta] = \int_0^1 \theta(1)d\theta = 1/2 \ .$$

(5) Unconditional Variance (or total variance) is

$$Var[X_t] = VHM + EPV = 1/12 + 1/2 = 7/12 \ .$$

1.1 Bühlmann Model

The Bühlmann model assumes that for any selected risk, the random variables { X_1, X_2, ..., X_N, X_{N+1}, …} are **independently and identically distributed**. For the selected risk, each X_t has the same probability distribution for any time period t, both for the X_1, X_2, ..., X_N random variables in the experience period, and future outcomes X_{N+1}, X_{N+2}, …. As Hans Bühlmann described it, "homogeneity *in time*" is assumed.

The characteristics that determine the risk's exposure to loss are assumed to be unchanging and the risk parameter θ associated with the risk is constant through time for the risk. The means and variances of the random variables for the different time periods are equal and are labeled $\mu(\theta)$ and $\sigma^2(\theta)$, respectively, as shown in the table below:

Assumptions of Bühlmann Credibility

Hypothetical Mean: $\mu(\theta) = E_{X|\Theta}[X_1|\theta] = \ldots = E_{X|\Theta}[X_N|\theta] = E_{X|\Theta}[X_{N+1}|\theta] = \ldots$

Process Variance: $\sigma^2(\theta) = Var_{X|\Theta}[X_1|\theta] = \ldots = Var_{X|\Theta}[X_N|\theta] = Var_{X|\Theta}[X_{N+1}|\theta] = \ldots$

Of course the hypothetical means and process variances will vary among risks, but they are assumed to be unchanging for any individual risk in the Bühlmann model.

To apply Bühlmann credibility, the average values of these quantities over the whole population of risks are needed, along with the variance of the hypothetical means for the population:

(1) Population mean: $\mu = E_{\Theta}[\mu(\Theta)] = E_{\Theta}[E_{X|\Theta}[X_t|\Theta]]$

(2) Expected Value of Process Variance: $EPV = E_{\Theta}[\sigma^2(\Theta)] = E_{\Theta}[Var_{X|\Theta}[X_t|\Theta]]$

(3) Variance of Hypothetical Means: $VHM = Var_{\Theta}[\mu(\Theta)] = E_{\Theta}[(\mu(\Theta) - \mu)^2]$.

The population mean $\mu = E_{\Theta}[E_{X|\Theta}[X_t|\Theta]]$ provides an estimate for the expected value of X_t in the absence of any prior information about the risk. The *EPV* indicates the variability to be expected from observations made about individual risks. The *VHM* is a measure of the differences in the means among risks in the population.

Because $\mu(\theta)$ is unknown for the selected risk, the mean $\bar{X} = \left(\frac{1}{N}\right)\sum_{t=1}^{N} X_t$ is used in the estimation process. It is an unbiased estimator for $\mu(\theta)$,

$$E_{X|\Theta}[\bar{X}\,|\theta] = E_{X|\Theta}[\left(\frac{1}{N}\right)\sum_{t=1}^{N} X_t\ |\theta] = \left(\frac{1}{N}\right)\sum_{t=1}^{N} E_{X|\Theta}[X_t|\theta] = \left(\frac{1}{N}\right)\sum_{t=1}^{N}\mu(\theta) = \mu(\theta)\ .$$

The conditional variance of $\bar{X}$, assuming independence of the X_t given θ, is

$$Var_{X|\Theta}[\bar{X}\,|\theta] = Var_{X|\Theta}[\left(\frac{1}{N}\right)\sum_{t=1}^{N} X_t\ |\theta] = \left(\frac{1}{N}\right)^2 \sum_{t=1}^{N} Var_{X|\Theta}[X_t|\theta] = \left(\frac{1}{N}\right)^2 \sum_{t=1}^{N}\sigma^2(\theta) = \frac{\sigma^2(\theta)}{N}\ .$$

The unconditional variance of $\bar{X}$ is

$$Var[\bar{X}] = Var_{\Theta}[E_{X|\Theta}[\bar{X}|\Theta]] + E_{\Theta}[Var_{X|\Theta}[\bar{X}|\Theta]] = Var_{\Theta}[\mu(\Theta)] + \frac{E_{\Theta}[\sigma^2(\Theta)]}{N} = VHM + \frac{EPV}{N}\,.$$

Bühlmann credibility assigned to estimator $\bar{X}$ is given by the well-known formula

$$Z = \frac{N}{N+K}\ ,$$

where *N* is the number of observations for the risk and *K* = *EPV* / *VHM*. Multiplying the numerator and denominator by (*VHM* / *N*) gives an alternative form:

$$Z = \frac{VHM}{VHM + \frac{EPV}{N}}\ .$$

Note that the denominator is just $Var[\bar{X}]$ as derived a few lines earlier. Therefore *Z* = *N* / (*N*+*K*) can be written as

$$Z = \frac{\text{Variance of the Hypothetical Means}}{\text{Total Variance of the Estimator } \bar{X}} = \frac{Var_{\Theta}[\mu(\Theta)]}{Var[\bar{X}]}\ .$$

The numerator is a measure of how far apart the means of the risks in the population are, while the denominator is a measure of the total variance of the estimator.

The credibility weighted estimate for $\mu(\theta) = E_{X|\Theta}[X_t\,|\,\theta]$, for t =1, 2, ..., N, N+1, ... is

$$\hat{\mu}(\theta) = Z \cdot \overline{X} + (1 - Z) \cdot \mu \quad .$$

The estimator $\hat{\mu}(\theta)$ is a linear least squares estimator for $\mu(\theta)$. This means that

$$E[\{[Z \cdot \overline{X} + (1 - Z) \cdot \mu] - \mu(\Theta)\}^2]$$

is minimized when $Z = N / (N+K)$. Appendix A proves this.[5]

1.2 Bühlmann-Straub Model

The requirement that the random variables $X_1, X_2, ..., X_N, X_{N+1}, \ldots$ for a risk be identically distributed is easily violated in the real world. For example:

- The work force of a workers compensation policyholder may change in size from one year to the next.
- The number of vehicles owned by a commercial automobile policyholder may change through time.
- The amount of earned premium for a rating class varies from year to year.

In all of these cases, one should not assume that $X_1, X_2, ..., X_N, X_{N+1}, \ldots$ are identically distributed, although an assumption of independence may be warranted.

A risk's exposure to loss may vary and it is assumed that this exposure can be measured. Some measures of exposure to loss are:

- Amount of insurance premium
- Number of employees
- Payroll
- Number of insured vehicles
- Number of claims

In fact, a fundamental premise of insurance rating is that exposure bases can be identified that are directly related to the potential for loss.

The Bühlmann-Straub model assumes that the means of the random variables are equal for the selected risk, but that the process variances are inversely proportional to the size (i.e., exposure) of the risk during each observation period. For example, when the risk is twice as large, the process variance is halved. These assumptions are summarized in the following table:

[5] The expected squared error is minimized only if the true values of the *EPV* and *VHM* are used to calculate *K*. If estimated values of the *EPV* and *VHM* (or *K*) are used, which is commonly done in practice, the linear estimator as given above is no longer optimal. This is an advanced topic beyond the scope of this study note.

Assumptions of Bühlmann-Straub Credibility

	Period 1	. . .	Period N	. . .
Exposure	m_1	. . .	m_N	. . .
Hypothetical Mean for Risk θ per Unit of Exposure	$\mu(\theta) = E_{X\mid\Theta}[X_1 \mid \theta] = \ldots$		$= E_{X\mid\Theta}[X_N \mid \theta] = \ldots$	
Process Variance for Risk θ	$Var_{X\mid\Theta}[X_1 \mid \theta] = \dfrac{\sigma^2(\theta)}{m_1}$	. . .	$Var_{X\mid\Theta}[X_N \mid \theta] = \dfrac{\sigma^2(\theta)}{m_N}$	. . .

The random variables X_t represent number of claims, monetary losses, or some other quantity of interest per unit of exposure, and m_t is the measure of exposure. For example, X_t could be number of claims per house-year[6], or X_t might be a loss ratio.[7] Note that the process variance for the random variable decreases as the exposure increases.

Example The annual numbers of claims for truck drivers in a homogeneous population are independently and identically distributed. [The population might represent the work force of a large trucking company with strict hiring standards and good safety training for each driver.] For each driver the number of claims per year has a mean of $\mu(\theta)$ and a variance of $\sigma^2(\theta)$. (The θ parameter applies to every driver in the group.)

A group of 10 drivers is selected from the larger population. (1) What is the expected annual claims frequency for the group of 10 drivers? (2) What is the variance of the annual claims frequency for the group?

Solution Let $X_{1t}, X_{2t}, \ldots, X_{10t}$ be random variables representing the number of claims in year t for each of the ten selected drivers. Then, $X_t = \left(\frac{1}{10}\right)\sum_{i=1}^{10} X_{it}$ is the annual claims frequency for the group; that is, it is the annual number of claims per driver. The exposure is $m_t = 10$ and the unit of exposure is one driver. The expected value and variance for the annual claims frequency for the group are

$$E_{X|\Theta}[X_t \mid \theta] = E_{X|\Theta}\left[\left(\frac{1}{10}\right)\sum_{i=1}^{10} X_{it} \mid \theta\right] = \left(\frac{1}{10}\right)\sum_{i=1}^{10} E_{X|\Theta}[X_{it} \mid \theta] = \left(\frac{1}{10}\right)\sum_{i=1}^{10}\mu(\theta) = \mu(\theta) \text{ and}$$

[6] A house-year means one house insured for one full year. It also represents two houses each insured for one-half year, or n houses each insured for $(1/n)$ years.
[7] Loss ratio equals losses divided by premium. In this case premium is the measure of exposure. A loss ratio of 60% means that there are .60 in losses for each 1.00 of premium.

$$Var_{X|\Theta}[X_t \mid \theta] = Var_{X|\Theta}\left[\left(\frac{1}{10}\right)\sum_{i=1}^{10} X_{it} \mid \theta\right] = \left(\frac{1}{10}\right)^2 \sum_{i=1}^{10} Var_{X|\Theta}[X_{it} \mid \theta] = \left(\frac{1}{100}\right)\sum_{i=1}^{10}\sigma^2(\theta) = \frac{\sigma^2(\theta)}{10}.$$

In this example, the exposure is the number of drivers in the group, which is 10. The expected claims frequency is the same whether there is one driver, 10 drivers, or 100 drivers in the group; however, the variance in the group's claims frequency is inversely proportional to the number of drivers in the group.

How should random variables $X_1, X_2, \ldots, X_N$ associated with a selected risk (or group of risks) be combined to estimate the hypothetical mean $\mu(\theta)$? A weighted average using the exposures m_t will give a linear estimator for $\mu(\theta)$ with minimum variance. First define

$$m = \sum_{t=1}^{N} m_t \ .$$

Then, define the weighted average

$$\overline{X} = \sum_{t=1}^{N}\left(\frac{m_t}{m}\right) X_t \ .$$

Recall that the variance of each X_t given θ is $\sigma^2(\theta) / m_t$. For a weighted average $\overline{X} = \sum_{t=1}^{N} w_t X_t$, the variance of $\overline{X}$ will be minimized by choosing the weights w_t to be inversely proportional to the variances of the individual X_t's; that is, random variables with smaller variances should be given more weight. So, weights $w_t = m_t / m$ are called for under the current assumptions. The proof is included as an exercise.

The conditional expected value and variance of $\overline{X}$ given risk parameter θ are

$$E_{X|\Theta}[\overline{X} \mid \theta] = E_{X|\Theta}\left[\sum_{t=1}^{N}\left(\frac{m_t}{m}\right) X_t \mid \theta\right] = \sum_{t=1}^{N}\left(\frac{m_t}{m}\right) E_{X|\Theta}[X_t \mid \theta] = \sum_{t=1}^{N}\left(\frac{m_t}{m}\right)\mu(\theta) = \mu(\theta) \text{ , and}$$

$$Var_{X|\Theta}[\overline{X} \mid \theta] = Var_{X|\Theta}\left[\sum_{t=1}^{N}\left(\frac{m_t}{m}\right) X_t \mid \theta\right] = \sum_{t=1}^{N}\left(\frac{m_t}{m}\right)^2 Var_{X|\Theta}[X_t \mid \theta] = \sum_{t=1}^{N}\left(\frac{m_t}{m}\right)^2 \left(\frac{\sigma^2(\theta)}{m_t}\right) = \frac{\sigma^2(\theta)}{m} \ .$$

<u>Example</u> Continuing the prior example, assume that the number of drivers in the group was six in the first year, seven in the second year and nine in the third year. X_t represents the number of claims <u>per</u> <u>driver</u> and m_t is the number of drivers in the group in years t = 1, 2, and 3.

(1) $m = 6 + 7 + 9 = 22$

(2) $\overline{X} = \dfrac{6X_1 + 7X_2 + 9X_3}{22}$

(3)

$$E_{X|\Theta}[\overline{X}|\,\theta] = E_{X|\Theta}\left[\frac{6X_1 + 7X_2 + 9X_3}{22}\,\middle|\,\theta\right] = \frac{6E_{X|\Theta}[X_1\,|\,\theta] + 7E_{X|\Theta}[X_2\,|\,\theta] + 9E_{X|\Theta}[X_3\,|\,\theta]}{22} = \mu(\theta)$$

(4)

$$Var_{X|\Theta}[\overline{X}|\,\theta] = Var_{X|\Theta}\left[\frac{6X_1 + 7X_2 + 9X_3}{22}\,\middle|\,\theta\right] = \frac{6^2 Var_{X|\Theta}[X_1\,|\,\theta] + 7^2 Var_{X|\Theta}[X_2\,|\,\theta] + 9^2 Var_{X|\Theta}[X_3\,|\,\theta]}{22^2}$$

$$= \frac{6^2\left(\dfrac{\sigma^2(\theta)}{6}\right) + 7^2\left(\dfrac{\sigma^2(\theta)}{7}\right) + 9^2\left(\dfrac{\sigma^2(\theta)}{9}\right)}{22^2} = \frac{\sigma^2(\theta)}{22} \qquad \blacksquare$$

Example A class for workers compensation insurance produced the following:

Year	**Payroll in 100 Units**	**Losses**	**Loss per Exposure**
1	100,000 = m_1	300,000	3.00 = x_1
2	110,000 = m_2	200,000	1.82 = x_2
3	120,000 = m_3	320,000	2.67 = x_3
Total	330,000 = m	820,000	2.48 = $\bar{x}$

The exposure unit is 100 of payroll. Note that $\bar{x}$ can be calculated two equivalent ways:

(1) $\bar{x} = \sum_{t=1}^{3}\left(\dfrac{m_t}{m}\right)x_t = \left(\dfrac{100{,}000}{330{,}000}\right)(3.00) + \left(\dfrac{110{,}000}{330{,}000}\right)(1.82) + \left(\dfrac{120{,}000}{330{,}000}\right)(2.67)$ [8]

(2) $\bar{x}$ = (300,000 + 200,000 + 320,000) / 330,000 . $\blacksquare$

The *EPV* and *VHM* are defined to be

$$EPV = E_{\Theta}[\sigma^2(\Theta)] \quad \text{and} \quad VHM = Var_{\Theta}[\mu(\Theta)]\ ,$$

where the expected value is over all risk parameters θ in the population. Remember, the loss per unit of exposure is used because the exposure can vary through time and from risk to risk.

The unconditional mean and variance of $\overline{X}$ are

[8] Note that this method produces 2.49, which differs from 2.48 in the table because of rounding error.

$$E[\overline{X}] = E_{\Theta}[E_{X|\Theta}[\overline{X}\big|\Theta]] = E_{\Theta}[\mu(\Theta)] = \mu \text{, and}$$

$$Var[\overline{X}] = Var_{\Theta}[E_{X|\Theta}[\overline{X}\big|\Theta]] + E_{\Theta}[Var_{X|\Theta}[\overline{X}\big|\Theta]] = Var_{\Theta}[\mu(\Theta)] + \frac{E_{\Theta}[\sigma^2(\Theta)]}{m} = VHM + \frac{EPV}{m}.$$

As in the simpler Bühlmann case, the credibility assigned to the estimator $\overline{X}$ of $\mu(\theta)$ is

$$Z = \frac{Variance\ \ of\ \ the\ \ Hypothetical\ \ Means}{Total\ \ Variance\ \ of\ \ the\ \ Estimator\ \ \overline{X}}$$

$$= \frac{VHM}{VHM + \frac{EPV}{m}} \quad .$$

Multiplying the numerator and denominator by (m/VHM) yields a familiar looking formula

$$Z = \frac{m}{m+K} \quad .$$

The total exposure m replaces N in the Bühlmann formula and the parameter K is defined as usual

$$K = \frac{EPV}{VHM} = \frac{E_{\Theta}[\sigma^2(\Theta)]}{Var_{\Theta}[\mu(\Theta)]} \quad .$$

Note that the Bühlmann model is actually a special case of the more general Bühlmann-Straub model with $m_t = 1$ for all t.

The credibility weighted estimate is

$$\hat{\mu}(\theta) = Z \cdot \overline{X} + (1 - Z) \cdot \mu \quad .$$

Example The actuaries at the Good Health Insurance Company calculate prospective premiums for group insurance policies using a Bühlmann-Straub credibility model. Analysis of Good Health's data led to the following assumptions for its business:

- For all policies together, the prospective average annual expected pure premium per insured person is 2,400.
- The variance of the hypothetical mean pure premiums across group plans is 500,000.
- The expected value of the process variance in annual costs per insured person is 250,000,000.

One of Good Health's clients had the following experience during a one-year period with costs adjusted to reflect prospective costs.

Group Policy	Insured Persons	Cost per Insured Person
1	240	3,000

Calculate a credibility weighted pure premium for group policy 1.

Solution K = (250,000,000/500,000) = 500, μ = 2,400, m = 240

$$\bar{X} = 3{,}000 \quad , \quad Z = \frac{240}{240+500} = .3243$$

Estimated Pure Premium = .3243(3,000) + (1 − .3243)(2,400) = 2,594.58 ▌▌

2. Estimation of Credibility Formula Parameters

Both the Bühlmann and Bühlmann-Straub models require a determination of the parameter K. In practice there are several ways this is done:

(1) Judgmentally select K. A larger K gives less credibility to the individual sample mean $\bar{X}$ and more credibility to the population mean. A smaller K gives more credibility to the individual sample mean $\bar{X}$, but the sample mean $\bar{X}$ may change significantly from one measurement period to another producing a fluctuating estimate. For example, the latest three years of data may be used to calculate $\bar{X}$. As an old year rolls off to be replaced by a more recent year of data, the value of $\bar{X}$ can drastically change.

(2) Select a K value that would have worked best in prior applications of the model. If one is trying to estimate $E[X_{N+1} \mid X_1, ..., X_N]$ from $Z\bar{X} + (1-Z)\mu$, one approach would be to minimize $\sum[x_{N+1} - (Z\bar{x} + (1-Z)\mu)]^2$ where the actual outcomes for year N+1 are compared with the credibility weighted forecasts for all risks in the population. One could also use some other function of the difference between outcomes and forecasts such as the sum of absolute errors or the sum of absolute values of the percentage errors.

(3) Attempt to determine the *EPV* and *VHM* components of K.

For the remainder of this study note, (3) will be discussed.

To calculate the Expected Value of the Process Variance (*EPV*) and the Variance of the Hypothetical Means (*VHM*) the following are needed:

(1) Process variances $\sigma^2(\theta)$ for each risk in the population
(2) Hypothetical means $\mu(\theta)$ for each risk in the population
(3) Distribution function for θ to calculate the $EPV = E_{\Theta}[\sigma^2(\Theta)]$ and $VHM = Var_{\Theta}[\mu(\Theta)]$

Either all of the above must be estimated from the data, or else simplifying assumptions must be made.

Suppose that there are R independent risks to be observed for N separate time periods as represented by the random variables in the table below:

Risk	**Time Period** 1	2	. . .	N	**Risk Parameter**
1	X_{11}	X_{12}	. . .	X_{1N}	θ_1
2	X_{21}	X_{22}	. . .	X_{2N}	θ_2
:	:	:	. . .	:	:
R	X_{R1}	X_{R2}	. . .	X_{RN}	θ_R

Note that another subscript has been added to the random variables because the discussion now concerns multiple risks. The X_{it}'s are random variables representing the losses or number of claims for risk i during time period t and, for now, each risk i has N independent outcomes. Associated with each θ_i is a mean $\mu(\theta_i)$ such that

$$E_{X\mid\Theta}[\, X_{it} \mid \theta_i \,] = \mu(\theta_i) \;.$$

The expected values of each of the N outcomes for any selected risk i are assumed to be equal.

The following additional assumptions are also made:

(1) For any selected risk i, the X_{it}'s are independent given $\Theta_i = \theta_i$.

(2) The outcomes for any risk are independent of any other risk.

(3) The random variables $\Theta_1, \Theta_2, ..., \Theta_R$ are independent and identically distributed from a common distribution $f_{\Theta}(\theta)$.

For a random variable X_{it} selected at random from the table the expected value is

$$E[X_{it}] = E_{\Theta}[E_{X\mid\Theta}[\, X_{it} \mid \Theta_i \,]\,] = E_{\Theta}[\mu(\Theta_i)] = \mu \;.$$

To calculate the unconditional variance of a randomly selected X_{it}, the total variance formula must be used:

$$Var[X_{it}] = Var_{\Theta}[E_{X\mid\Theta}[\, X_{it} \mid \Theta_i \,]] + E_{\Theta}[Var_{X\mid\Theta}[\, X_{it} \mid \Theta_i \,]] \,.$$

2.1 Nonparametric Estimation

In the nonparametric case, no assumptions are made about the form or parameters of the distributions of X_{it}, nor are any assumptions made about the distribution of the risk parameters Θ_i. Of course, for the Bühlmann model it is assumed that for any given risk i, the random variables $\{ X_{i1}, X_{i2}, ..., X_{iN}\}$, representing the outcomes for N different observations, are independently and identically distributed with identical means and variances. The outcomes for different risks are also independent. In the Bühlmann-Straub model, the N_i outcomes for risk i have the same means but the process variances are inversely related to the exposure. Note that the number of observations N_i has a subscript indicating that the number of observations can vary by risk in the Bühlmann-Straub model.

To apply the models in Sections 1.1 and 1.2, some information about the probability distributions $f_{X\mid\Theta}(x \mid \Theta = \theta)$ and $f_{\Theta}(\theta)$ is required. Although the exact distribution functions are not needed, the *EPV* and *VHM* must be obtained in order to calculate K and then the credibility Z. In practice the *EPV* and *VHM* are often unknown. The *EPV* and *VHM* can be estimated from the data for the Bühlmann model or the more complicated Bühlmann-Straub model. The estimation procedures are sometimes referred to as empirical Bayesian procedures or, equivalently, empirical Bayes estimation.

<u>Bühlmann Model</u>

Estimates of the *EPV* and *VHM* can be made from empirical observations about a sample from the population of risks. Assume that there are R risks in the sample and N separate observations will be made for each risk. The $R\cdot N$ random variables in the left-hand section of the following table represent the outcomes:

	Time Period				**Risk's Sample Mean**	**Risk's Sample**
Risk	**1**	**2**	**. . .**	**N**	$\bar{X}_i$	**Process Variance**
1	X_{11}	X_{12}	. . .	X_{1N}	$\bar{X}_1 = \left(\frac{1}{N}\right)\sum_{t=1}^{N} X_{1t}$	$\hat{\sigma}_1^2 = \left(\frac{1}{N-1}\right)\sum_{t=1}^{N}(X_{1t} - \bar{X}_1)^2$
2	X_{21}	X_{22}	. . .	X_{2N}	$\bar{X}_2 = \left(\frac{1}{N}\right)\sum_{t=1}^{N} X_{2t}$	$\hat{\sigma}_2^2 = \left(\frac{1}{N-1}\right)\sum_{t=1}^{N}(X_{2t} - \bar{X}_2)^2$
:	:	:	. . .	:	:	:
R	X_{R1}	X_{R2}	. . .	X_{RN}	$\bar{X}_R = \left(\frac{1}{N}\right)\sum_{t=1}^{N} X_{Rt}$	$\hat{\sigma}_R^2 = \left(\frac{1}{N-1}\right)\sum_{t=1}^{N}(X_{Rt} - \bar{X}_R)^2$

The $\overline{X}_i$'s are unbiased estimators for each risk's mean $\mu(\theta_i)$, and the $\hat{\sigma}_i^2$'s are unbiased estimators for each risk's process variance $\sigma^2(\theta_i)$. Note that the divisor is $(N-1)$ rather than N for the sample variances because sample means $\overline{X}_i$ rather than true means $\mu(\theta_i)$ are used to compute them.

Expected Value of the Process Variance – *EPV*

The individual sample process variances $\hat{\sigma}_i^2$ can be combined to produce an unbiased estimate for the expected value of the process variance of the population. To estimate this, the average of the individual sample process variances is computed:

$$E\hat{P}V = \left(\frac{1}{R}\right)\sum_{i=1}^{R}\hat{\sigma}_i^2 = \left(\frac{1}{R(N-1)}\sum_{i=1}^{R}\sum_{t=1}^{N}(X_{it}-\overline{X}_i)^2\right) .$$

Variance of the Hypothetical Means – *VHM*

In the prior table the $\overline{X}_i$'s are estimators for the unknown hypothetical means $\mu(\theta_i)$ and these values can be used to estimate the variance of the hypothetical means. Because the R risks are all independent, the $\overline{X}_i$'s are independent random variables. An unbiased estimator for the variance of $\overline{X}_i$ is

$$V\hat{a}r[\overline{X}_i] = \left(\frac{1}{R-1}\right)\sum_{i=1}^{R}(\overline{X}_i - \overline{X})^2 \quad \text{where} \quad \overline{X} = \left(\frac{1}{R}\sum_{i=1}^{R}\overline{X}_i\right) .$$

However, this is NOT the estimate for the *VHM*.

The total variance formula gives

$$Var[\overline{X}_i] = Var_{\Theta}[E_{X\mid\Theta}[\overline{X}_i \mid \Theta_i]] + E_{\Theta}[Var_{X\mid\Theta}[\overline{X}_i \mid \Theta_i]] .$$

The first term on the right can be simplified by noting that $E_{X\mid\Theta}(\overline{X}_i \mid \Theta_i = \theta_i) = \mu(\theta_i)$. The following relationship can simplify the second term on the right:

$$Var_{X|\Theta}\left[\overline{X}_i \,|\, \theta_i\right] = Var_{X|\Theta}\left[\left(\frac{1}{N}\right)\sum_{t=1}^{N}X_{it} \,|\, \theta_i\right] = \left(\frac{1}{N}\right)^2\sum_{t=1}^{N}Var_{X|\Theta}[X_{it} \,|\, \theta_i] = \sigma^2(\theta_i)/N .$$

Substituting into the formula

$$Var[\overline{X}_i] = Var_{\Theta}[\mu(\Theta_i)] + E_{\Theta}[\sigma^2(\Theta_i)]/N .$$

The first term on the right is just the *VHM* and the second term is the *EPV* / *N*. Rearranging terms yields

$$VHM = Var[\bar{X}_i] - EPV / N .$$

Substituting in the unbiased estimators for the quantities on the right produces an unbiased estimator for the *VHM*,

$$V\hat{H}M = \left(\frac{1}{R-1}\right)\sum_{i=1}^{R}(\bar{X}_i - \bar{X})^2 - \left\{\left(\frac{1}{R(N-1)}\right)\sum_{i=1}^{R}\sum_{t=1}^{N}(X_{it} - \bar{X}_i)^2\right\}/ N .$$

To estimate *VHM* = $Var_{\Theta}[\mu(\Theta_i)]$, the quantity $V\hat{a}r[\bar{X}_i]$ needs to be adjusted downward by $E\hat{P}V / N$ because the process variance increases the variability of the estimates $\bar{X}_i$ for $\mu(\theta_i)$. The larger the average process variance, the larger the necessary correction to $V\hat{a}r[\bar{X}_i]$ to estimate *VHM* = $Var_{\Theta}[\mu(\Theta_i)]$. Because of this subtraction, it is also possible that the estimator $V\hat{H}M$ may be negative.

What does it mean for a *VHM* to be negative? Because variances must be nonnegative, one usually concludes that zero is a reasonable estimate for the *VHM* and that the means of the individual risks are all the same. There is no empirical evidence that the risk means are different from one another.

<u>Example</u> Two risks were selected at random from a population. Risk 1 had 0 claims in year one, 3 claims in year two, and 0 claims in year three: (0,3,0). The claims by year for Risk 2 were (2,1,2). In this case, $R = 2$ and $N = 3$.

$\bar{x}_1 = (0+3+0)/3 = 1$, $\bar{x}_2 = (2+1+2)/3 = 5/3$, and $\bar{x} = [1+(5/3)]/2 = 4/3$
$\hat{\sigma}_1^2 = [(0-1)^2 + (3-1)^2 + (0-1)^2]/(3-1) = 3$
$\hat{\sigma}_2^2 = [(2-5/3)^2 + (1-5/3)^2 + (2-5/3)^2]/(3-1) = 1/3$

$$E\hat{P}V = (\hat{\sigma}_1^2 + \hat{\sigma}_2^2)/2 = [3+(1/3)]/2 = 5/3$$

$$V\hat{H}M = \left(\frac{1}{2-1}\right)\left\{[1-(4/3)]^2 + [(5/3)-(4/3)]^2\right\} - (5/3)/3 = -1/3$$

The sample means for the two risks are close relative to the sizes of the sample process variances. The calculated $V\hat{H}M$ is negative, so a value of zero will be assumed. The hypothetical means are indistinguishable. This implies a credibility factor $\hat{Z} = 0$. ▌▌

Credibility weighted estimators for the risk means can be derived using the formulas

$$\hat{K} = \frac{E\hat{P}V}{V\hat{H}M}, \quad \hat{Z} = \frac{N}{N + \hat{K}} \quad \text{and} \quad \hat{\mu}(\theta_i) = \hat{Z} \cdot \overline{X}_i + (1 - \hat{Z}) \cdot \overline{X} .$$

Although the estimators $E\hat{P}V$ and $V\hat{H}M$ are unbiased estimators for the Expected Value of the Process Variance and the Variance of the Hypothetical Means, the estimated value $\hat{Z}$ for the credibility Z is <u>not</u> unbiased. In practice, the above $\hat{Z}$ is generally accepted as a reasonable estimate for the credibility weight.

<u>Example</u> Two risks were selected at random from a population. Over a four-year period, Risk 1 had the following claims by year: (0,0,1,0). The claims by year for Risk 2 were: (2,1,0,2) . Calculate credibility weighted estimates for the expected number of claims per year for each risk.

<u>Solution</u> $\bar{x}_1 = (0+0+1+0)/4 = 1/4$, $\bar{x}_2 = (2+1+0+2)/4 = 5/4$, and
$\bar{x} = [(1/4)+(5/4)]/2 = 3/4$

$$\hat{\sigma}_1^2 = [(0-(1/4))^2 + (0-(1/4))^2 + (1-(1/4))^2 + (0-(1/4))^2]/(4-1) = 1/4$$
$$\hat{\sigma}_2^2 = [(2-(5/4))^2 + (1-(5/4))^2 + (0-(5/4))^2 + (2-(5/4))^2]/(4-1) = 11/12$$

$$E\hat{P}V = (\hat{\sigma}_1^2 + \hat{\sigma}_1^2)/2 = [(1/4)+(11/12)]/2 = 7/12$$

$$V\hat{H}M = \left(\frac{1}{2-1}\right)\left\{[(1/4)-(3/4)]^2 + [(5/4)-(3/4)]^2\right\} - (7/12)/4 = 17/48$$

$$\hat{K} = \frac{E\hat{P}V}{V\hat{H}M} = \frac{7/12}{17/48} = 28/17$$

$$\hat{Z} = \frac{N}{N+\hat{K}} = \frac{4}{4+(28/17)} = 17/24$$

$$\hat{\mu}(\theta_1) = \hat{Z} \cdot \overline{X}_1 + (1-\hat{Z}) \cdot \overline{X} = (17/24)\cdot(1/4) + (7/24)\cdot(3/4) = 19/48 = .3958$$

$$\hat{\mu}(\theta_2) = \hat{Z} \cdot \overline{X}_2 + (1-\hat{Z}) \cdot \overline{X} = (17/24)\cdot(5/4) + (7/24)\cdot(3/4) = 53/48 = 1.1042$$

.

<u>Bühlmann-Straub Model</u>

The Bühlmann-Straub Model is more complicated because a risk's exposure to loss can vary from year to year, and the number of years of observations can change from risk to risk. The reason that Bühlmann-Straub can handle varying numbers of years is because the number of years of data for a risk is reflected in the total exposure for the risk.

The table below shows X_{it} − representing claim frequency, loss ratio, or average cost per exposure − in the top half of the cell, and the corresponding number of exposures for the risk during the same time period in the bottom half:

Risk	**Periods of Experience**					
1	X_{11}	X_{12}	$\cdots$	X_{1N_1}		
	m_{11}	m_{12}	$\cdots$	m_{1N_1}		
2	X_{21}	X_{22}	$\cdots$	$\cdots$	$\cdots$	X_{2N_2}
	m_{21}	m_{22}	$\cdots$	$\cdots$	$\cdots$	m_{2N_2}
:	:	:	$\cdots$	$\cdots$	$\cdots$	:
R		X_{R1}	X_{R2}	$\cdots$	$\cdots$	X_{RN_R}
		m_{R1}	m_{R2}	$\cdots$	$\cdots$	m_{RN_R}

The number of periods of experience can vary by risk; for example, for Risk 2 there are N_2 periods of experience compared with N_1 periods for Risk 1. The experience periods do not have to start at the same time either. For example, the first experience period for Risk 1 might be in Year Y whereas the first experience period for Risk R may be Year Y+1.

Example ABC Insurance, Inc. sells dental insurance plans to companies with fewer than one hundred employees. An actuary is analyzing the number of claims per employee. Looking at the first company in her file, she sees that the company has three full years of plan coverage. In the first year there were 40 employee-years with 84 claims, in the second year there were 44 employee-years with 88 claims, and in the third year there were 42 employee-years with 105 claims.

Designating this selected company as Risk 1, then:

X_{11} = 84 claims / 40 employee-years = 2.1 claims/employee-year
X_{12} = 88 claims / 44 employee-years = 2.0 claims/employee-year
X_{13} = 105 claims / 42 employee-years = 2.5 claims/employee-year

The exposures are m_{11} = 40 employee-years, m_{12} = 44 employee-years, and m_{13} = 42 employee-years.

The next table shows estimators for risk means and variances:

Risk	Periods	Total Exposure	Sample Mean	Sample Estimate for σ^2
1	N_1	$m_1 = \sum_{t=1}^{N_1} m_{1t}$	$\bar{X}_1 = \left(\frac{1}{m_1}\right)\sum_{t=1}^{N_1} m_{1t} X_{1t}$	$\hat{\sigma}_1^2 = \left(\frac{1}{N_1 - 1}\right)\sum_{t=1}^{N_1} m_{1t}(X_{1t} - \bar{X}_1)^2$
2	N_2	$m_2 = \sum_{t=1}^{N_2} m_{2t}$	$\bar{X}_2 = \left(\frac{1}{m_2}\right)\sum_{t=1}^{N_2} m_{2t} X_{2t}$	$\hat{\sigma}_2^2 = \left(\frac{1}{N_2 - 1}\right)\sum_{t=1}^{N_2} m_{2t}(X_{2t} - \bar{X}_2)^2$
:	:	:	:	:
R	N_R	$m_R = \sum_{t=1}^{N_R} m_{Rt}$	$\bar{X}_R = \left(\frac{1}{m_R}\right)\sum_{t=1}^{N_R} m_{Rt} X_{Rt}$	$\hat{\sigma}_R^2 = \left(\frac{1}{N_R - 1}\right)\sum_{t=1}^{N_R} m_{Rt}(X_{Rt} - \bar{X}_R)^2$
Total		$m = \sum_{i=1}^{R} m_i$	$\bar{X} = \sum_{i=1}^{R} m_i \bar{X}_i / m$	$E\hat{P}V = \sum_{i=1}^{R}(N_i - 1)\hat{\sigma}_i^2 / \left(\sum_{i=1}^{R}(N_i - 1)\right)$

In the Bühlmann-Straub model, the mean is assumed to be constant through time for each risk i:

$$\mu(\theta_i) = E_{X\mid\Theta}[X_{i1} \mid \theta_i] = E_{X\mid\Theta}[X_{i2} \mid \theta_i] = \ldots = E_{X\mid\Theta}[X_{iN_i} \mid \theta_i] \ .$$

$\bar{X}_i$ is an unbiased estimator for the mean of risk i:

$$E_{X|\Theta}[\bar{X}_i \mid \theta_i] = E_{X|\Theta}\left[\left(\frac{1}{m_i}\right)\sum_{t=1}^{N_i} m_{it} X_{it} \mid \theta_i\right] = \left(\frac{1}{m_i}\right)\sum_{t=1}^{N_i} m_{it} E_{X|\Theta}\left[X_{it} \mid \theta_i\right] = \left(\frac{1}{m_i}\right)\sum_{t=1}^{N_i} m_{it}\mu(\theta_i) = \mu(\theta_i) \ .$$

Recall from Section 1.2 that the process variance of X_{it} is inversely proportional to the exposure: $Var_{X\mid\Theta}[X_{it} \mid \theta_i] = \sigma^2(\theta_i) / m_{it}$. This means that for risk i,

$$\sigma^2(\theta_i) = m_{it} Var_{X\mid\Theta}[X_{it} \mid \theta_i] \text{ , for } t = 1 \text{ to } N_i \ .$$

To provide some motivation for the process variance estimates in the last column of the table above, it is helpful to write out the definition of variance as

$$\sigma^2(\theta_i) = m_{it} E_{X|\Theta}[(X_{it} - \mu(\theta_i))^2 \mid \theta_i] \text{, for } t = 1 \text{ to } N_i \ .$$

Summing both sides over t and dividing by the number of terms N_i yields

$$\left(\frac{1}{N_i}\right)\sum_{t=1}^{N_i} \sigma^2(\theta_i) = \left(\frac{1}{N_i}\right)\sum_{t=1}^{N_i} m_{it} E_{X|\Theta}[(X_{it} - \mu(\theta_i))^2 \mid \theta_i] \text{ , or}$$

$$\sigma^2(\theta_i) = E_{X|\Theta}\left[\left(\frac{1}{N_i}\right)\sum_{t=1}^{N_i} m_{it}(X_{it} - \mu(\theta_i))^2 \middle| \theta_i\right].$$

The quantity $\mu(\theta_i)$ is unknown, so $\bar{X}_i$ is used instead in the estimation process. This reduces the degrees of freedom by one so N_i is replaced by $(N_i - 1)$ in the denominator:

$$\sigma^2(\theta_i) = E_{X|\Theta}\left[\left(\frac{1}{N_i - 1}\right)\sum_{t=1}^{N_i} m_{it}(X_{it} - \bar{X}_i)^2 \mid \theta_i\right].$$

Thus, an unbiased[9] estimator for $\sigma^2(\theta_i)$ is

$$\hat{\sigma}_i^2 = \left(\frac{1}{N_i - 1}\right)\sum_{t=1}^{N_i} m_{it}(X_{it} - \bar{X}_i)^2 .$$

If $m_{it} \equiv 1$ and each risk has the same number of years of data, then the estimators $\bar{X}_i$ and $\hat{\sigma}_i^2$ match those from the Bühlmann model.

<u>Expected Value of the Process Variance</u>

The *EPV* can be estimated by combining process variance estimates $\hat{\sigma}_i^2$ of the R risks. If they are combined with weights $w_i = (N_i - 1)/\left(\sum_{i=1}^{R}(N_i - 1)\right)$, then

$$E\hat{P}V = \sum_{i=1}^{R} w_i \hat{\sigma}_i^2 = \left(\frac{\sum_{i=1}^{R}\sum_{t=1}^{N_i} m_{it}(X_{it} - \bar{X}_i)^2}{\sum_{i=1}^{R}(N_i - 1)}\right).$$

This is an unbiased estimator for the *EPV* as shown in Appendix C. A way to "guess" the denominator is to observe that there are $\sum_{i=1}^{R} N_i$ terms added together in the numerator, an indication of degrees of freedom. However, there are R estimators $\bar{X}_i$ for the individual

[9] A proof is not very difficult and can be modeled after the standard proof that the sample variance with divisor (N−1) is an unbiased estimator for the variance.

means, and that reduces the degrees of freedom by R. So, the divisor is $\sum_{i=1}^{R} N_i - R =$ $\sum_{i=1}^{R}(N_i - 1)$.

Variance of the Hypothetical Means

The hypothetical mean for risk i is $\mu(\theta_i)$. The variance of the hypothetical means can be written as

$$VHM = E_{\Theta}[(\mu(\Theta_i) - \mu)^2] \quad \text{where} \quad \mu = E_{\Theta}[\mu(\Theta_i)] .$$

Because outcomes of the random variables $\overline{X}_i$ and $\overline{X}$ are estimators for $\mu(\theta_i)$ and μ, respectively, a good starting point for developing an estimator for the variance of the hypothetical means is: $\frac{1}{(R-1)}\sum_{i=1}^{R} m_i(\overline{X}_i - \overline{X})^2$. Note that each term is weighted by its total exposure over the experience period. However, this is not unbiased. An unbiased estimator is

$$V\hat{H}M = \left(\sum_{i=1}^{R} m_i(\overline{X}_i - \overline{X})^2 - (R-1)E\hat{P}V\right) / \left(m - \left(\frac{1}{m}\right)\sum_{i=1}^{R} m_i^2\right) .$$

Showing that this estimator is unbiased takes several steps and the details are included in Appendix C.

With the Bühlmann-Straub model, the measure to use in the credibility formula is the total exposure for risk i over the whole experience period. (Note that the inputs in the $E\hat{P}V$ and $V\hat{H}M$ formulas are measured per unit of exposure.) The formulas to compute credibility weighted estimates are

$$\hat{K} = \frac{E\hat{P}V}{V\hat{H}M}, \qquad \hat{Z}_i = \frac{m_i}{m_i + K} \qquad \text{and} \qquad \hat{\mu}(\theta_i) = \hat{Z}_i \cdot \overline{X}_i + (1 - \hat{Z}_i) \cdot \overline{X} .$$

Example Carpentry contractors A and B had insurance policies covering pickup trucks. Over a four-year period the following was observed:

		Year			
Insured		***Y***	***Y*+1**	***Y*+2**	***Y*+3**
A	Number of Claims	3	2	2	0
	Insured Vehicles	2	2	2	1
B	Number of Claims	2	1	0	
	Insured Vehicles	4	3	2	

Estimate the expected annual claim frequency for each insured using the Bühlmann-Straub model.

Solution The random variables X_{it} representing claim frequency and the corresponding exposures m_{it} are as follows:

		Year			
Insured		***Y***	***Y*+1**	***Y*+2**	***Y*+3**
A	Claims per Exposure	$X_{11} = 3/2$	$X_{12} = 1$	$X_{13} = 1$	$X_{14} = 0$
	Exposure = Number of Vehicles	$m_{11} = 2$	$m_{12} = 2$	$m_{13} = 2$	$m_{14} = 1$
B	Claims per Exposure	$X_{21} = 1/2$	$X_{22} = 1/3$	$X_{23} = 0$	
	Exposure = Number of Vehicles	$m_{21} = 4$	$m_{22} = 3$	$m_{23} = 2$	

The claim frequency X_{it} is (Number of Claims)/(Insured Vehicles). The first table shows number of claims that are the values for ($m_{it}X_{it}$).

$m_A = 2 + 2 + 2 + 1 = 7$, $m_B = 4 + 3 + 2 = 9$, and $m = 7 + 9 = 16$.

$\bar{x}_A = (3+2+2+0)/7 = 1$, $\bar{x}_B = (2+1+0)/9 = 1/3$, and
$\bar{x} = [(7)(1)+(9)(1/3)]/16 = 5/8$

$$\hat{\sigma}_A^2 = [2((3/2)-1)^2 + 2(1-1)^2 + 2(1-1)^2 + 1(0-1)^2]/(4-1) = 1/2$$

$$\hat{\sigma}_B^2 = [4((1/2)-(1/3))^2 + 3((1/3)-(1/3))^2 + 2(0-(1/3))^2]/(3-1) = 1/6$$

$$E\hat{P}V = \frac{(4-1)(1/2)+(3-1)(1/6)}{(4-1)+(3-1)} = 11/30 = .3667$$

Note: You can sum the seven terms inside the brackets in the expressions for $\hat{\sigma}_A^2$ and $\hat{\sigma}_B^2$ *and then divide the total by [(4−1)+(3−1)]. This saves the step of dividing by the individual degrees of freedom and then undoing it when calculating the EPV estimate.*

$$V\hat{H}M = \frac{7[1-(5/8)]^2 + 9[(1/3)-(5/8)]^2 - (2-1)(11/30)}{16 - \left(\frac{1}{16}\right)(7^2 + 9^2)} = .1757$$

$$\hat{K} = \frac{E\hat{P}V}{V\hat{H}M} = .3667 / .1757 = 2.0871$$

$$\hat{Z}_{\mathrm{A}} = \frac{m_{\mathrm{A}}}{m_{\mathrm{A}} + \hat{K}} = \frac{7}{7 + 2.0871} = .7703 \quad \rightarrow \quad \hat{\mu}_{\mathrm{A}} = .7703(1) + (1 - .7703)(5/8) = .9139$$

$$\hat{Z}_{\mathrm{B}} = \frac{m_{\mathrm{B}}}{m_{\mathrm{B}} + \hat{K}} = \frac{9}{9 + 2.0871} = .8118 \quad \rightarrow \quad \hat{\mu}_{\mathrm{B}} = .8118(1/3) + (1 - .8118)(5/8) = .3882 \quad \blacksquare$$

<u>Balancing the estimators</u>

It is desirable in many cases for the estimators $\hat{\mu}(\theta_i) = \hat{Z}_i \overline{X}_i + (1 - \hat{Z}_i)\overline{X}$, when weighted together, to equal the overall sample mean $\overline{X} = \sum_{i=1}^{R} \frac{m_i}{m} \overline{X}_i$. An example is an experience rating plan. The amount of premium charged to pay for losses per unit of exposure for the i^{th} risk is $\hat{\mu}(\theta_i)$. The average loss per exposure for all risks is $\overline{X}$. To make the experience rating plan balanced, i.e., for the sum of the pieces to add up to the total, the goal is

$$\overline{X} = \sum_{i=1}^{R} \frac{m_i}{m} [\hat{Z}_i \overline{X}_i + (1 - \hat{Z}_i)\overline{X}] \ .$$

In general this might not happen. Putting $\hat{\mu}$ in for the complement of credibility yields

$$\overline{X} = \sum_{i=1}^{R} \frac{m_i}{m} \overline{X}_i = \sum_{i=1}^{R} \frac{m_i}{m} \left(\hat{Z}_i \overline{X}_i + (1 - \hat{Z}_i)\hat{\mu} \right)$$

$$\sum_{i=1}^{R} m_i \overline{X}_i = \sum_{i=1}^{R} m_i \hat{Z}_i \overline{X}_i + \sum_{i=1}^{R} m_i (1 - \hat{Z}_i)\hat{\mu}$$

$$\sum_{i=1}^{R} m_i \overline{X}_i - \sum_{i=1}^{R} m_i \hat{Z}_i \overline{X}_i = \hat{\mu} \sum_{i=1}^{R} m_i (1 - \hat{Z}_i)$$

$$\frac{\sum_{i=1}^{R} m_i (1 - \hat{Z}_i)\overline{X}_i}{\sum_{i=1}^{R} m_i (1 - \hat{Z}_i)} = \hat{\mu} \quad .$$

The left-hand side can be simplified by noting that

$$m_i (1 - \hat{Z}_i) = m_i \left(1 - \frac{m_i}{m_i + \hat{K}} \right) = \frac{m_i \hat{K}}{m_i + \hat{K}} = \hat{K}\hat{Z}_i \, .$$

The weighted average of the credibility estimates for the R risks will equal $\overline{X}$ if

$$\frac{\sum_{i=1}^{R}\hat{K}\hat{Z}_i\overline{X}_i}{\sum_{i=1}^{R}\hat{K}\hat{Z}_i} = \frac{\sum_{i=1}^{R}\hat{Z}_i\overline{X}_i}{\sum_{i=1}^{R}\hat{Z}_i} = \hat{\mu} \quad .$$

Balance can be achieved by using a credibility weighted $\hat{\mu}$ as the complement of credibility.

<u>Example</u> The prior example produced:

Weights	**Credibilities**	**Sample Means**	**Credibility Estimates**
$\frac{m_{\mathrm{A}}}{m} = \frac{7}{16}$	$\hat{Z}_{\mathrm{A}} = .7703$	$\overline{X}_{\mathrm{A}} = 1$	$\hat{\mu}_{\mathrm{A}} = .9139$
$\frac{m_{\mathrm{B}}}{m} = \frac{9}{16}$	$\hat{Z}_{\mathrm{B}} = .8118$	$\overline{X}_{\mathrm{B}} = \frac{1}{3}$	$\hat{\mu}_{\mathrm{B}} = .3882$
		$\overline{X} = \frac{5}{8} = .6250$	Weighted Average = .6182

The overall sample mean is .6250 but the weighted average of the credibility estimates is .6182, which is 1.1% below the overall sample mean.

The credibility weighted average is

$$\hat{\mu} = \frac{\sum_{i=1}^{2}\hat{Z}_i\overline{X}_i}{\sum_{i=1}^{2}\hat{Z}_i} = \frac{.7703(1) + .8118(1/3)}{.7703 + .8118} = .6579 \ .$$

Recalculating the credibility estimates: $\hat{\mu}_{\mathrm{A}} = .7703(1) + (1 - .7703)(.6579) = .9214$ and $\hat{\mu}_{\mathrm{B}} = .8118(1/3) + (1 - .8118)(.6579) = .3944$.

The weighted average is (7/16)(.9214) + (9/16)(.3944) = .6250, which equals $\overline{X}$. ▌▌

2.2 Semiparametric Estimation

Assuming that the random variables X_{it} have a particular distributional form can simplify the calculations. For example, the probability distribution for X_{it} might be Poisson or binomial.

If X_{it} is the number of <u>claims</u> <u>per</u> <u>exposure</u> and m_{it} is the number of exposures for risk i, then the product $m_{it}X_{it}$ is the number of claims for risk i in time period t.

Furthermore, assume that the number of claims is <u>Poisson</u> distributed. If the risk parameter θ_i is the mean number of claims per exposure, then $m_{it}\theta_i$ is both the mean and variance of the number of claims for exposure m_{it},

$$m_{it}\theta_i = E_{X\mid\Theta}[m_{it}X_{it} \mid \theta_i] = Var_{X\mid\Theta}[m_{it}X_{it} \mid \theta_i] = m_{it}^2\, Var_{X\mid\Theta}[X_{it} \mid \theta_i] \,.$$

Dividing through by m_{it} yields

$$E_{X\mid\Theta}[X_{it} \mid \theta_i] = m_{it} Var_{X\mid\Theta}[X_{it} \mid \theta_i] = \theta_i \,.$$

By definition, $E_{X\mid\Theta}[X_{it} \mid \theta_i] = \mu(\theta_i)$ and $m_{it}Var_{X\mid\Theta}[X_{it} \mid \theta_i] = \sigma^2(\theta_i)$. The result is that with the Poisson assumption, it follows that $\mu(\theta_i) = \sigma^2(\theta_i)$. Because the mean and process variance are equal for each risk, the same is true for the expected values with

$$E_\Theta[\mu(\Theta_i)] = E_\Theta[\sigma^2(\Theta_i)] = EPV \,.$$

The estimator for the mean $\mu = E_\Theta[\mu(\Theta_i)]$ is $\hat{\mu} = \overline{X}$, and the prior formula means that this same estimator can be used for the EPV under the Poisson assumption.

<u>Example</u> The information in the prior example will be used along with the additional assumption that the number of claims for each risk is Poisson distributed.

As calculated previously, $\bar{x} = [(7)(1) + (9)(1/3)]/16 = 5/8$. This is the estimate for the overall mean, so under the Poisson assumption it follows that

$$E\hat{P}V = \bar{x} = 5/8 = .625 \,.$$

Without the Poisson assumption, the EPV estimate was .3667 . Continuing on with the calculations,

$$V\hat{H}M = \frac{7[1-(5/8)]^2 + 9[(1/3)-(5/8)]^2 - (2-1)(5/8)}{16 - \left(\frac{1}{16}\right)(7^2+9^2)} = .1429$$

$$\hat{K} = \frac{E\hat{P}V}{V\hat{H}M} = \frac{5/8}{.1429} = 4.3737$$

$$\hat{Z}_{\mathrm{A}} = \frac{m_{\mathrm{A}}}{m_{\mathrm{A}} + \hat{K}} = \frac{7}{7+4.3737} = .6155 \quad \rightarrow \quad \hat{\mu}_{\mathrm{A}} = .6155(1) + (1-.6155)(5/8) = .8558$$

$$\hat{Z}_{\mathrm{B}} = \frac{m_{\mathrm{B}}}{m_{\mathrm{B}} + \hat{K}} = \frac{9}{9+4.3737} = .6730 \quad \rightarrow \quad \hat{\mu}_{\mathrm{B}} = .6730(1/3) + (1-.6730)(5/8) = .4287 \,.$$

Of course, one could use the balancing procedure if balanced estimates were appropriate.

In the example above, one can compute the $E\hat{P}V$ either from the sample process variances of the data or with a Poisson assumption. Unlike the prior example, the next example requires an assumption about the claims process to estimate the $E\hat{P}V$.

Example During a three-year period, a group of 1,000 auto policies generated the following claims profile:

Total Number of Claims in Three Years	**Number of Policies**
0	533
1	320
2	105
3	22
4	12
5	8

Each policy was in the group for the entire three years and insured exactly one automobile. The expected number of claims per year for each insured is assumed to be constant from year to year and the actual number of claims per year follows a Poisson distribution.

Determine a credibility estimator for the expected annual claims frequency for a policy that had no claims over the entire three-year period. Do the same for a policy that had five claims.

Solution Without the Poisson assumption, additional information would be needed to solve this problem. To estimate the *EPV*, it would be necessary to know the number of claims during each of the three years for each of the 1,000 policies. This would allow the calculation of a sample process variance for each insured, and then an overall average process variance could be computed.

If each insured has a Poisson distribution, then as stated previously

$$E_{\Theta}[\mu(\Theta_i)] = E_{\Theta}[\sigma^2(\Theta_i)] = EPV\,,$$

and $\hat{\mu} = \overline{X}$ can be used as an estimator for the *EPV* which produces

$$E\hat{P}V = \overline{X} = \frac{533(0)+320(1)+105(2)+22(3)+12(4)+8(5)}{(1{,}000)(3)} = .2280\ .$$

The 3 in the denominator is required because $\overline{X}$ is the average annual claims frequency.

The simpler Bühlmann model can be used because all of the exposures are identically one so that $m_{it} \equiv 1$, and

$$V\hat{H}M = Var(\overline{X}_i) - E\hat{P}V/N \quad \rightarrow \quad V\hat{H}M = \left(\frac{1}{R-1}\right)\sum_{i=1}^{R}(\overline{X}_i - \overline{X})^2 - E\hat{P}V/N \ .$$

Note that R = 1,000 and that there are 533 terms in the summation with $\overline{X}_i = 0$, 320 terms with $\overline{X}_i = 1/3$, 105 terms with $\overline{X}_i = 2/3$, etc. The threes in the denominators are required because the annual frequency is the number of claims over the three-year period divided by three.

$$V\hat{H}M = \frac{533(0-.228)^2 + 320(1/3-.228)^2 + 105(2/3-.228)^2 + 22(3/3-.228)^2 + 12(4/3-.228)^2 + 8(5/3-.228)^2}{(1000-1)} - .228/3$$

$$V\hat{H}M = .0199, \qquad \hat{K} = \frac{E\hat{P}V}{V\hat{H}M} = \frac{.2280}{.0199} = 11.46, \qquad \hat{Z} = \frac{N}{N+\hat{K}} = \frac{3}{3+11.46} = .2075$$

$$\hat{\mu}_{0\text{ claims}} = .2075(0) + (1-.2075)(.2280) = .1807$$

$$\hat{\mu}_{5\text{ claims}} = .2075(5/3) + (1-.2075)(.2280) = .5265 \ \blacksquare$$

3. Conclusion

The “Credibility” chapter of *Foundations of Casualty Actuarial Science* along with this study note provides a basic education in credibility theory. As with most academic presentations, many of the examples are idealized and do not address some “messy” realities that make it difficult to estimate credibility model parameters. In practice, using the most precise credibility parameter estimate, versus a reasonable estimate, should not affect final results that much. Mahler discusses this in [5].

Credibility theory produces a linear least-squares estimator and this can sometimes be a major source of error.[10] For example, suppose that for some line of insurance, class rates vary from a low of 0.10 per exposure to as much as 100.00 per exposure. An error of 2.00 in the class rate would not be so bad for the 100.00 rate but it would be a huge error for a class that deserved a 0.10 rate. In the Bühlmann and Bühlmann-Straub models, it is the size of the error that matters so that a 2.00 error in the 0.10 rate gets the same weight as a 2.00 error in the 100.00 rate. Rather than minimizing the squared errors as Bühlmann and Bühlmann-Straub credibility do, it may be preferable to minimize the squared relative errors, i.e. percentage errors. One way to accomplish this is to take logarithms of values and then minimize squared errors. Errors in logs are relative errors in the original scale. Another alternative is to use class loss ratios instead of pure premiums.

[10] Gary Venter provided this example.

It can be difficult to identify a good complement of credibility,[11]the quantity that gets multiplied by $(1 - Z)$. In the basic Bühlmann and Bühlmann-Straub models, this quantity is the population or class mean to which the risk belongs. What are the consequences if this population or class mean is also highly variable and indicated insurance rates fluctuate significantly? The selection of a good complement of credibility is sometimes part art and part science. Boor identifies criteria to consider when choosing a complement of credibility in [6].

The models covered in this study note assume that a random variable, representing an important quantity for a risk, has a constant mean $\mu(\theta)$ through time and that the variance $\sigma^2(\theta)$ is also constant. In reality, a risk's characteristics may shift through time. This situation is addressed by Mahler in [7], [8], and [9].

Risk heterogeneity is another practical issue that must be considered. A big risk is not necessarily the sum of smaller independent risks. For example, a risk with 100,000 in annual premium may not behave like the sum of ten independent 10,000 premium risks, even though the risks come from the same rating classification. Suppose the actuary is looking at loss ratios. Often, the variance in the loss ratio for the 100,000 risk will be greater than the variance in the loss ratio for the sum of ten independent 10,000 risks. A larger risk can have different risk characteristics. The larger risk should receive less credibility than implied by the simple $m/(m+K)$ formula. Mahler addresses this in [9].

[11] Also called "complement for credibility."

<u>Appendix A</u>: Bühlmann and Bühlmann-Straub Credibility Estimators are Linear Least Squares Estimators

A risk is selected at random from a population and N observations are to be made. These N outcomes are represented by the random variables $\{X_1, X_2, \ldots, X_N\}$. Each outcome has the same mean $E_{X \mid \Theta}[X_t \mid \theta] = \mu(\theta)$ where θ is the risk parameter associated with the selected risk.

In the Bühlmann model the random variables are independently and identically distributed and the sample mean is given by

$$\overline{X} = \left(\frac{1}{N}\right)\sum_{t=1}^{N} X_t \ .$$

In the Bühlmann-Straub model, the random variables all have the same mean $\mu(\theta)$, but the conditional variances (conditional on risk parameter θ) are inversely proportional to the exposure m_t, which can vary from observation to observation. In this case the sample mean is defined to be

$$\overline{X} = \sum_{t=1}^{N}\left(\frac{m_t}{m}\right)X_t \ , \quad \text{where } \ m = \sum_{t=1}^{N} m_t \ .$$

It follows that $E_{X \mid \Theta}[\ \overline{X} \mid \theta\ ] = \mu(\theta)$ for both models. If the expectation is calculated over the entire population with possibly different risk parameters, then

$$E[\overline{X}] = E_{\Theta}[E_{X \mid \Theta}[X \mid \Theta]] = E_{\Theta}[\mu(\Theta)] = \mu \ ,$$

where μ is the population mean.

The outcome of the random variable $\overline{X}$ will be used to estimate $\mu(\theta)$. In particular, the goal is to find a linear estimator $a + b\overline{X}$ with the two constants selected to

$$\textbf{Minimize } \ E[(a + b\overline{X} - \mu(\Theta))^2] \ .$$

The expectation is over $\overline{X}$ and Θ. If $\overline{X}$ were the linear least squares estimator for $\mu(\Theta)$, then $a = 0$ and $b = 1$ would minimize the above expected value.

Minimizing the expectation will start with a rearrangement of terms:

$$\begin{aligned} E[(a + b\overline{X} - \mu(\Theta))^2] &= E[(a + b\overline{X} - b\mu(\Theta) + b\mu(\Theta) - \mu(\Theta))^2] \\ &= E[\left(b(\overline{X} - \mu(\Theta)) + a - (1-b)\mu(\Theta)\right)^2] \end{aligned}$$

$$= E[b^2(\bar{X}-\mu(\Theta))^2 + 2b(\bar{X}-\mu(\Theta))(a-(1-b)\mu(\Theta)) + (a-(1-b)\mu(\Theta))^2]$$

$$= E[b^2(\bar{X}-\mu(\Theta))^2] + 2E[b\{\bar{X}-\mu(\Theta)\}\{a-(1-b)\mu(\Theta)\}] + E[\{a-(1-b)\mu(\Theta)\}^2] \ .$$

The middle term is zero because

$$E[b\{\bar{X}-\mu(\Theta)\}\{a-(1-b)\mu(\Theta)\}] = E_\Theta[E_{X|\Theta}[b\{a-(1-b)\mu(\Theta)\}\{\bar{X}-\mu(\Theta)\}|\Theta]]$$

$$= E_\Theta[b\{a-(1-b)\mu(\Theta)\}E_{X|\Theta}[\{\bar{X}-\mu(\Theta)\}|\Theta]] = 0 \ .$$

Although the above looks ugly, the point is that $E_{X|\Theta}[\{\bar{X}-\mu(\Theta)\}|\Theta] = 0$ because the expected value of $\bar{X}$ conditional on Θ is $\mu(\Theta)$.

So far the result is

$$E[(a+b\bar{X}-\mu(\Theta))^2] = b^2E[\{\bar{X}-\mu(\Theta)\}^2] + E[\{a-(1-b)\mu(\Theta)\}^2] \ .$$

Only the second term on the right involves a. What value of a will minimize it?

$$E[\{a-(1-b)\mu(\Theta)\}^2] = E[a^2 - 2a(1-b)\mu(\Theta) + (1-b)^2\mu^2(\Theta)]$$

$$= a^2 - 2a(1-b)E[\mu(\Theta)] + (1-b)^2E[\mu^2(\Theta)]$$

Taking the partial derivative with respect to a of the right-hand side and setting it equal to zero, and replacing $E[\mu(\Theta)]$ by the population mean μ, yields

$$a = (1-b)\mu \ .$$

Substituting this expression in for a:

$$E[(a+b\bar{X}-\mu(\Theta))^2] = b^2E[\{\bar{X}-\mu(\Theta)\}^2] + (1-b)^2E[\{\mu(\Theta)-\mu\}^2] .$$

The first term following b^2 is EPV/N for the Bühlmann model or EPV/m for the Bühlmann-Straub model. The following relationships show this for the Bühlmann model:

$$E[\{\bar{X}-\mu(\Theta)\}^2] = E_\Theta[E_{X|\Theta}[\{\bar{X}-\mu(\Theta)\}^2|\Theta]] = E_\Theta[Var_{X|\Theta}[\bar{X}|\Theta]] = E_\Theta[Var_{X|\Theta}[X_t|\Theta]]/N \ .$$

The term following $(1-b)^2$ is the VHM, so for the Bühlmann model it follows that

$$E[(a+b\overline{X}-\mu(\Theta))^2] = b^2(EPV/N) + (1-b)^2 VHM \; .$$

Minimizing the expected value on the left is equivalent to minimizing the right-hand side. The derivative of the right-hand side with respect to *b* is

$$2b(EPV/N) - 2(1-b)VHM = 0,$$

which yields

$$b = \frac{VHM}{VHM + \dfrac{EPV}{N}} \; .$$

This can be rewritten as $b = \dfrac{N}{N + \dfrac{EPV}{VHM}} = \dfrac{N}{N+K}$, $K = \dfrac{EPV}{VHM}$.

If the expressions calculated above for the two constants are substituted into the estimator $a + b\overline{X}$, then

$$a + b\overline{X} = \left(\frac{N}{N+K}\right)\overline{X} + \left(1 - \frac{N}{N+K}\right)\mu \; .$$

This is Bühlmann’s form. For the Bühlmann-Straub model, the *N* is replaced by the total exposure *m*.

<u>Appendix B</u>: Proof of Unconditional (Total) Variance Formula

(1) <u>Total Variance Formula</u>:

Let W represent $E_{X|\Theta}[X \mid \Theta]$; then the formula $Var_{\Theta}[W] = E_{\Theta}[W^2] - (E_{\Theta}[W])^2$ leads immediately to

$$Var_{\Theta}[E_{X|\Theta}[X \mid \Theta]] = E_{\Theta}[\{E_{X|\Theta}[X \mid \Theta]\}^2] - \{E_{\Theta}[E_{X|\Theta}[X \mid \Theta]]\}^2 . \qquad (1)$$

Similarly, $Var_{X|\Theta}[X \mid \Theta] = E_{X|\Theta}[X^2 \mid \Theta] - \{E_{X|\Theta}[X \mid \Theta]\}^2$ yields

$$E_{\Theta}[Var_{X|\Theta}[X \mid \Theta]] = E_{\Theta}[E_{X|\Theta}[X^2 \mid \Theta]] - E_{\Theta}[\{E_{X|\Theta}[X \mid \Theta]\}^2] . \qquad (2)$$

Note that in equations (1) and (2) there is a common term on the right-hand sides, though of opposite signs, $E_{\Theta}[\{E_{X|\Theta}[X \mid \Theta]\}^2]$.

Adding (1) and (2) together yields

$$Var_{\Theta}[E_{X|\Theta}[X \mid \Theta]] + E_{\Theta}[Var_{X|\Theta}[X \mid \Theta]] = E_{\Theta}[E_{X|\Theta}[X^2 \mid \Theta]] - \{E_{\Theta}[E_{X|\Theta}[X \mid \Theta]]\}^2 .$$

The expectations on the right-hand side can be rewritten as

$$E_{\Theta}[E_{X|\Theta}[X^2 \mid \Theta]] - \{E_{\Theta}[E_{X|\Theta}[X \mid \Theta]]\}^2 = E[X^2] - \{E[X]\}^2 = Var[X],$$

proving that

$$Var[X] = Var_{\Theta}[E_{X|\Theta}[X \mid \Theta]] + E_{\Theta}[Var_{X|\Theta}[X \mid \Theta]] .$$

Appendix C: Nonparametric Estimators for the Expected Value of the Process Variance and the Variance of the Hypothetical Means in the Bühlmann-Straub Model are Unbiased

Assumptions:

(1) Each of R independent risks has an associated risk parameter θ_i.

(2) For the i^{th} risk there are N_i observation periods and the random variable X_{it} represents the observation for risk i in time period t.

(3) The number of exposures for the i^{th} risk in time period t is m_{it} and the sum for N_i periods is $m_i = \sum_{t=1}^{N_i} m_{it}$. The average of the X_{it} over N_i observation periods is defined to be $\overline{X}_i = \sum_{t=1}^{N_i} \frac{m_{it}}{m_i} X_{it}$. The mean for all risks is $\overline{X} = \sum_{i=1}^{R} \frac{m_i}{m} \overline{X}_i$ with $m = \sum_{i=1}^{R} m_i$.

(4) The expected value of X_{it} is constant through time for a given risk with risk parameter θ_i:

$$E_{X\,|\,\Theta}[X_{it} \mid \theta_i] = \mu(\theta_i) \text{ for } t = 1 \text{ to } N_i .$$

(5) The variance of X_{it} for a given risk with risk parameter θ_i is inversely proportional to the amount of exposure:

$$Var_{X\,|\,\Theta}[X_{it} \mid \theta_i] = \sigma^2(\theta_i) / m_{it} \text{ for } t = 1 \text{ to } N_i .$$

Expected Value of the Process Variance

The expected value of the estimator for the Expected Value of the Process Variance[12] is

$$E[E\hat{P}V] = E\left[\frac{\sum_{i=1}^{R}\sum_{t=1}^{N_i} m_{it}(X_{it} - \overline{X}_i)^2}{\sum_{i=1}^{R}(N_i - 1)}\right].$$

The innermost sum in the numerator can be rewritten as

[12] See Section 2.1 .

$$\sum_{t=1}^{N_i} m_{it}(X_{it}-\overline{X}_i)^2 = \sum_{t=1}^{N_i} m_{it}(X_{it}-\mu(\theta_i)+\mu(\theta_i)-\overline{X}_i)^2$$

$$= \sum_{t=1}^{N_i} m_{it}(X_{it}-\mu(\theta_i))^2 - 2\sum_{t=1}^{N_i} m_{it}(X_{it}-\mu(\theta_i))(\overline{X}_i-\mu(\theta_i)) + \sum_{t=1}^{N_i} m_{it}(\overline{X}_i-\mu(\theta_i))^2 .$$

The middle and last terms can be combined using $m_i = \sum_{t=1}^{N_i} m_{it}$ and $\overline{X}_i = \sum_{t=1}^{N_i} \frac{m_{it}}{m_i} X_{it}$, so

$$\sum_{t=1}^{N_i} m_{it}(X_{it}-\overline{X}_i)^2 = \left(\sum_{t=1}^{N_i} m_{it}(X_{it}-\mu(\theta_i))^2\right) - m_i(\overline{X}_i-\mu(\theta_i))^2 \ .$$

The conditional expectation of the first term on the right is

$$E_{X|\Theta}\left[\sum_{t=1}^{N_i} m_{it}(X_{it}-\mu(\theta_i))^2 \mid \theta_i\right] = \sum_{t=1}^{N_i} m_{it} Var_{X|\Theta}[X_{it} \mid \theta_i] = \sum_{t=1}^{N_i} m_{it}\left(\frac{\sigma^2(\theta_i)}{m_{it}}\right) = N_i\sigma^2(\theta_i) \ .$$

Because $\overline{X}_i = \sum_{t=1}^{N_i} \frac{m_{it}}{m_i} X_{it}$ and $E_{X|\Theta}[\overline{X}_i] = \mu(\theta_i)$, it follows that

$$E_{X|\Theta}[m_i(\overline{X}_i-\mu(\theta_i))^2 \mid \theta_i] = m_i Var_{X|\Theta}[\overline{X}_i \mid \theta_i] = m_i \sum_{t=1}^{N_i} \frac{m_{it}^2}{m_i^2} Var_{X|\Theta}[X_{it} \mid \theta_i]$$

$$= m_i \sum_{t=1}^{N_i} \frac{m_{it}^2}{m_i^2}\left(\frac{\sigma^2(\theta_i)}{m_{it}}\right) = \sigma^2(\theta_i)\sum_{t=1}^{N_i} \frac{m_{it}}{m_i} = \sigma^2(\theta_i) \ .$$

Combining the results gives

$$E_{X|\Theta}\left[\sum_{t=1}^{N_i} m_{it}(X_{it}-\overline{X}_i)^2 \middle| \theta_i\right] = N_i\sigma^2(\theta_i) - \sigma^2(\theta_i) = (N_i-1)\sigma^2(\theta_i) \ .$$

Substituting this back into the equation at the beginning of the section:

$$E[E\hat{P}V] = E\left[\frac{\sum_{i=1}^{R}\sum_{t=1}^{N_i} m_{it}(X_{it}-\overline{X}_i)^2}{\sum_{i=1}^{R}(N_i-1)}\right] = E_{\Theta}\left[\frac{\sum_{i=1}^{R} E_{X|\Theta}\left[\sum_{t=1}^{N_i} m_{it}(X_{it}-\overline{X}_i)^2 \middle| \Theta_i\right]}{\sum_{i=1}^{R}(N_i-1)}\right] = E_{\Theta}\left[\frac{\sum_{i=1}^{R}(N_i-1)\sigma^2(\Theta_i)}{\sum_{i=1}^{R}(N_i-1)}\right]$$

$$= \frac{\sum_{i=1}^{R}(N_i - 1)E_{\Theta}\left[\sigma^2(\Theta_i)\right]}{\sum_{i=1}^{R}(N_i - 1)} = \frac{\sum_{i=1}^{R}(N_i - 1)EPV}{\sum_{i=1}^{R}(N_i - 1)} = EPV .$$

This last equation shows that the estimator for the *EPV* is unbiased.

Variance of the Hypothetical Means

The expected value of the estimator for the Variance of the Hypothetical Means is

$$E[V\hat{H}M] = E\left[\left(\sum_{i=1}^{R} m_i(\overline{X}_i - \overline{X})^2 - (R-1)E\hat{P}V\right) / \left(m - \left(\frac{1}{m}\right)\sum_{i=1}^{R} m_i^2\right)\right] .$$

The summation in the numerator on the right-hand side can be rewritten as

$$\sum_{i=1}^{R} m_i(\overline{X}_i - \overline{X})^2 = \sum_{i=1}^{R} m_i(\overline{X}_i - \mu + \mu - \overline{X})^2$$

$$= \sum_{i=1}^{R} m_i(\overline{X}_i - \mu)^2 - 2\sum_{i=1}^{R} m_i(\overline{X}_i - \mu)(\overline{X} - \mu) + \sum_{i=1}^{R} m_i(\overline{X} - \mu)^2 .$$

The middle and last terms can be combined using $m = \sum_{i=1}^{R} m_i$ and $\overline{X} = \sum_{i=1}^{R} \frac{m_i}{m}\overline{X}_i$, so

$$\sum_{i=1}^{R} m_i(\overline{X}_i - \overline{X})^2 = \sum_{i=1}^{R} m_i(\overline{X}_i - \mu)^2 - m(\overline{X} - \mu)^2 .$$

Taking the expectation of both sides yields

$$E\left[\sum_{i=1}^{R} m_i(\overline{X}_i - \overline{X})^2\right] = \left(\sum_{i=1}^{R} m_i E\left[(\overline{X}_i - \mu)^2\right]\right) - mE\left[(\overline{X} - \mu)^2\right]$$

$$= \left(\sum_{i=1}^{R} m_i Var[\overline{X}_i]\right) - mVar[\overline{X}] .$$

The first variance on the right can be expanded as

$$Var[\overline{X}_i] = Var_{\Theta}[E_{X|\Theta}[\overline{X}_i \mid \Theta_i]] + E_{\Theta}[Var_{X|\Theta}[\overline{X}_i \mid \Theta_i]] .$$

Because $E_{X|\Theta}[\overline{X}_i \mid \theta_i] = \mu(\theta_i)$, the first term is the Variance of the Hypothetical Means, $Var_\Theta[\mu(\Theta_i)]$. Writing out the variance inside the second term yields

$$Var_{X|\Theta}[\overline{X}_i \mid \theta_i] = Var_{X|\Theta}\left(\sum_{t=1}^{N_i} \frac{m_{it} X_{it}}{m_i} \middle| \theta_i\right) = \sum_{t=1}^{N_i} \frac{m_{it}^2}{m_i^2} Var_{X|\Theta}[X_{it} \mid \theta_i] = \sum_{t=1}^{N_i} \frac{m_{it}^2}{m_i^2} \frac{\sigma^2(\theta_i)}{m_{it}} = \frac{\sigma^2(\theta_i)}{m_i} .$$

The result is

$$Var[\overline{X}_i] = Var_\Theta[\mu(\Theta_i)] + \frac{E_\Theta[\sigma^2(\Theta_i)]}{m_i} = VHM + EPV / m_i .$$

This result is also useful in calculating $Var[\overline{X}]$ as follows:

$$Var[\overline{X}] = Var\left[\sum_{i=1}^{R} \frac{m_i}{m} \overline{X}_i\right] = \sum_{i=1}^{R} \frac{m_i^2}{m^2} Var[\overline{X}_i] = \sum_{i=1}^{R} \frac{m_i^2}{m^2}\left(VHM + \frac{EPV}{m_i}\right)$$

$$= \left(\sum_{i=1}^{R} \frac{m_i^2}{m^2}\right) VHM + \frac{EPV}{m} .$$

Putting the pieces together yields

$$E\left[\sum_{i=1}^{R} m_i (\overline{X}_i - \overline{X})^2\right] = \left(\sum_{i=1}^{R} m_i \left(VHM + \frac{EPV}{m_i}\right)\right) - m\left(\left(\sum_{i=1}^{R} \frac{m_i^2}{m^2}\right) VHM + \frac{EPV}{m}\right).$$

Simplifying and combining terms gives

$$E\left[\sum_{i=1}^{R} m_i (\overline{X}_i - \overline{X})^2\right] = \left(m - \left(\frac{1}{m}\right) \sum_{i=1}^{R} m_i^2\right) VHM + (R-1)EPV .$$

Putting this result into the equation for the expected value of the estimator for the Variance of the Hypothetical Means yields

$$E\left[V\hat{H}M\right] = \left\{\left(m - \left(\frac{1}{m}\right) \sum_{i=1}^{R} m_i^2\right) VHM + (R-1)EPV - (R-1)E[E\hat{P}V]\right\} / \left(m - \left(\frac{1}{m}\right) \sum_{i=1}^{R} m_i^2\right).$$

The first section of this appendix showed $E[E\hat{P}V] = EPV$, leaving

$$E[V\hat{H}M] = VHM .$$

REFERENCES

Only a few selected references are listed here. The student new to credibility theory should find [2] to [4] particularly helpful. [2], [3] and [4] include extensive references.

[1] Bühlmann, Hans, "Experience Rating and Credibility," *ASTIN Bulletin*, Vol. 4, No. 3, 1967, pp. 199-207.
A classic paper that introduced the Bühlmann model.

[2] Herzog, Thomas N., *Introduction to Credibility Theory*, ACTEX Publications.
A well organized survey of credibility theory.

[3] Klugman, Stuart A., Harry H. Panjer and Gordon E. Willmot, *Loss Models: From Data to Decisions*, John Wiley & Sons.
Provides a more rigorous treatment of credibility theory and includes most the material covered in this study note.

[4] Mahler, Howard C. and Curtis Gary Dean, "Credibility," *Foundations of Casualty Actuarial Science*, Fourth Edition, Casualty Actuarial Society, 2001, Chapter 8.
An introduction to credibility theory with many examples and problems.

[5] Mahler, Howard C., "An Actuarial Note on Credibility Parameters," *Proceedings of the Casualty Actuarial Society*, Vol. LXXIII, 1986, pp.1-26.
Explains that reasonable, as opposed to precise, credibility formula parameters are generally sufficient.

[6] Boor, Joseph A., "The Complement of Credibility," *Proceedings of the Casualty Actuarial Society*, Vol. LXXXIII, 1996, pp. 1-40.
Discusses criteria to consider when choosing a complement of credibility.

[7] Mahler, Howard C., "An Example of Credibility and Shifting Risk Parameters," *Proceedings of the Casualty Actuarial Society*, Vol. LXXVII, 1990, pp. 225-308.
This reading and the next two consider situations where a risk's characteristics shift through time.

[8] Mahler, Howard C., "A Markov Chain Model of Shifting Risk Parameters," *Proceedings of the Casualty Actuarial Society*, Vol. LXXXIV, 1997, pp. 581-659.

[9] Mahler, Howard C., "Credibility with Shifting Risk Parameters, Risk Heterogeneity, and Parameter Uncertainty," *Proceedings of the Casualty Actuarial Society*, Vol. LXXXV, 1998, pp. 455-653.
As its title indicates, this paper extends credibility theory to several situations where simpler credibility models are not appropriate.

Exercises

1. Credibility Models

1. Random variables X and Y are independent and both have the same mean value: $E[X] = E[Y] = \mu$. The variances of X and Y are, respectively, $Var[X] = \sigma_X^2$ and $Var[Y] = \sigma_Y^2$. Define the random variable Z to be a linear combination of X and Y with $Z = w_1X + w_2Y$. Given that $E[Z] = \mu$, prove that $Var[Z]$ will be minimized if weights are selected such that $w_1 = \dfrac{(1/\sigma_X^2)}{(1/\sigma_X^2)+(1/\sigma_Y^2)}$ and $w_2 = \dfrac{(1/\sigma_Y^2)}{(1/\sigma_X^2)+(1/\sigma_Y^2)}$; that is, choose weights inversely proportional to the variances of each random variable.

2. Two urns contain a large number of balls with each ball marked with one number from the set {0,2,4}. The proportion of each type of ball in each urn is displayed in the table below:

Urn (denoted by *Θ*)	**Number on Ball**		
	0	**2**	**4**
A	60%	30%	10%
B	10%	30%	60%

An urn is randomly selected and then a ball is drawn at random from the urn. The number on the ball is represented by the random variable X.

(a) Calculate the hypothetical means (or conditional means) $E_{X\,|\,\Theta}[X \mid \Theta = \text{A}]$ and $E_{X\,|\,\Theta}[X \mid \Theta = \text{B}]$.
(b) Calculate the variance of the hypothetical means: $Var_{\Theta}[E_{X\,|\,\Theta}[X \mid \Theta]]$.
(c) Calculate the process variances (or conditional variances) $Var_{X\,|\,\Theta}[X \mid \Theta = \text{A}]$ and $Var_{X\,|\,\Theta}[X \mid \Theta = \text{B}]$.
(d) Calculate the expected value of the process variance: $E_{\Theta}[Var_{X\,|\,\Theta}[X \mid \Theta]]$.
(e) Calculate the total variance (or unconditional variance) $Var[X]$ and show that it equals the sum of the quantities calculated in (b) and (d).

1.1 Bühlmann Model

Many exercises are included in [4], the Mahler and Dean "Credibility" chapter of *Foundations of Casualty Actuarial Science.*

3. You are given:
   (i) Two risks have the following severity distributions:

Amount of Claim	Probability of Claim Amount for Risk 1	Probability of Claim Amount for Risk 2
250	0.5	0.7
2,500	0.3	0.2
60,000	0.2	0.1

(ii) Risk 1 is twice as likely to be observed as Risk 2.

A claim of 250 is observed.

Determine the Bühlmann credibility estimate of the second claim amount from the same risk. [Course 4 − Fall 2003 − #23]

1.2 Bühlmann-Straub Model

Use the following information for exercises 4-6.

Two urns contain a large number of balls with each ball marked with one number from the set {0,2,4}. Balls are drawn from the urns with replacement. The proportion of each type of ball in each urn is displayed in the table below:

Urn	**Number on Ball**		
	0	**2**	**4**
A	60%	30%	10%
B	10%	30%	60%

4. Suppose that urn A is selected and *n* balls are drawn from the urn.
   (a) What is the expected average value of the *n* balls?
   (b) What is the variance of the average value of the balls?

5. An urn is selected at random, two balls are drawn from the urn and the average value $\bar{X}$ of the two balls is recorded.
   (a) What is the expected value of the process variance (*EPV*) of $\bar{X}$?
   (b) What is the variance of the hypothetical means (*VHM*) of $\bar{X}$?

6. An urn is selected at random. During the first round, two balls were drawn and the average value of the two balls was 2.0. During the second round, four balls were drawn from the same urn as in the first round and the average of the four balls was 1.5.

   Another ball will be drawn from the same selected urn. Using the Bühlmann-Straub credibility model, what is the estmated value of the ball?

7. Bogus Advertising, Inc. has a small fleet of company cars that varies in size from year to year. All of the company drivers receive the same training and it is assumed that each car in the fleet has the same expected annual accident frequency that remains constant through time. This expected accident frequency per car is unknown but has a uniform distribution on the interval [0,1]. The number of claims for a car is Poisson distributed.

   During the last three years Bogus had the following claims experience:

Year	**Cars in Fleet**	**Total Number of Claims**
Y	4	1
Y+1	5	2
Y+2	2	0

   If Bogus expects to have three cars in the fleet next year, use Bühlmann-Straub credibility to estimate the total number of claims next year.

8. You are given four classes of insureds, each of whom may have zero or one claim, with the following probabilities:

Class	**Number of Claims**	
	0	**1**
I	0.9	0.1
II	0.8	0.2
III	0.5	0.5
IV	0.1	0.9

   A class is selected at random (with probability ¼), and four insureds are selected at random from the class. The total number of claims is two.

   If five insureds are selected at random from the same class, estimate the total number of claims using Bühlmann-Straub credibility. [Course 4 − Fall 2002 − #32]

9. You are given the following information on large business policyholders:

   i. Losses for each employee of a given policyholder are independent and have a common mean and variance.
   ii. The overall average loss per employee for all policyholders is 20.
   iii. The variance of the hypothetical means is 40.
   iv. The expected value of the process variance is 8,000.
   v. The following experience is observed for a randomly selected policyholder:

Year	Average Loss per Employee	Number of Employees
1	15	800
2	10	600
3	5	400

Determine the Bühlmann-Straub credibility premium per employee for this policyholder. [Course 4 – Fall 2001 − #26]

10. For each of the n independent random variables $X_1, X_2, ..., X_n$ the following are true:

$$E[X_i] = \mu \quad \text{and} \quad Var[X_i] = \sigma^2 / m_i \, .$$

The weighted mean is defined to be $\overline{X} = \sum_{i=1}^{n} w_i X_i$ with $\sum_{i=1}^{n} w_i = 1$.

(a) Prove that $E[\overline{X}] = \mu$.

(b) Prove that $Var[\overline{X}]$ is minimized by choosing weights $w_i = m_i / m$ where $m = \sum_{i=1}^{n} m_i$.

11. You are given:
    (i) The number of claims incurred in a month by any insured has a Poisson distribution with mean λ.
    (ii) The claim frequencies of different insureds are independent.
    (iii) The prior distribution is gamma with probability density function:

$$f(\lambda) = \frac{(100\lambda)^6 e^{-100\lambda}}{120\lambda} \quad .$$

(iv)

Month	Number of Insureds	Number of Claims
1	100	6
2	150	8
3	200	11
4	300	?

Determine the Bühlmann-Straub credibility estimate of the number of claims in Month 4. [Course 4 − Fall 2003 − #27]

<u>2.1 Nonparametric Estimation</u>

12. Two vehicles were selected at random from a population and the following claim counts were observed:

Vehicle	Number of Claims during Year			
	Year 1	Year 2	Year 3	Year 4
1	1	0	1	0
2	2	3	3	1

You are interested in the annual claims frequency of each vehicle. Use empirical Bayesian estimation procedures to do the following:

(a) Estimate the expected value of the process variance $E\hat{P}V$ for the number of claims in one year.

(b) Estimate the variance of the hypothetical means $V\hat{H}M$.

(c) Calculate the credibility weighted estimate of the annual claims frequency for each vehicle.

13. Two medium-sized insurance policies produced the following losses over a three-year period:

Insured	Annual Losses		
	Year 1	Year 2	Year 3
1	5	4	3
2	5	6	7

You are trying to estimate the expected annual losses for each insured. Assuming that the total exposures for each policy are equal and remain constant through time, use empirical Bayesian estimation procedures to do the following:

(a) Estimate the expected value of the process variance $E\hat{P}V$ for one year of losses.

(b) Estimate the variance of the hypothetical means $V\hat{H}M$.

(c) Calculate the credibility weighted estimate of the annual losses for each insured.

14. An insurer has data on losses for four policyholders for 7 years. The loss from the i^{th} policyholder for year j is X_{ij} .

You are given that

$$\sum_{i=1}^{4} \sum_{j=1}^{7} (X_{ij} - \overline{X}_i)^2 = 33.60$$

$$\sum_{i=1}^{4} (\overline{X}_i - \overline{X})^2 = 3.30 \ .$$

Using nonparametric empirical Bayes estimation, calculate the Bühlmann credibility factor for an individual policyholder. [Course 4 − Spring 2000 − #15 and Fall 2002 − #11]

15. XYZ Insurance Company offers a janitorial services policy that is rated on a per employee basis. The two insureds shown in the table below were randomly selected from XYZ's policyholder database. Over a four-year period the following was observed:

		Year			
Insured		***Y***	***Y*+1**	***Y*+2**	***Y*+3**
A	Number of Claims	3	2	3	1
	No. of Employees	2	2	2	1
B	Number of Claims	0	1	1	
	No. of Employees	4	4	4	

Estimate the expected annual claim frequency per employee for each insured using the empirical Bayes Bühlmann-Straub estimation model.

16. You are given the following information on towing losses for two classes of insureds - adults and youths:

Exposures

Year	Adult	Youth	Total
Y	2,000	450	2,450
Y+1	1,000	250	1,250
Y+2	1,000	175	1,175
Y+3	1,000	125	1,125
Total	5,000	1,000	6,000

Pure Premium

Year	Adult	Youth	Total
Y	0	15	2.755
Y+1	5	2	4.400
Y+2	6	15	7.340
Y+3	4	1	3.667
Total	3	10	4.167

You are also given that the estimated variance of the hypothetical means is 17.125 .

(a) Determine the nonparametric empirical Bayes credibility pure premium for the youth class, using $\hat{\mu} = \overline{X}$ as the complement of credibility.

(b) Determine the nonparametric empirical Bayes credibility pure premium for the youth class, using the method that preserves total losses. [Course 4 − Fall 2000 − #27]

17. You are given the following experience for two insured groups:

Group		Year			
		Y	***Y*+1**	***Y*+2**	**Total**
1	Number of members	8	12	5	25
	Average loss per member	96	91	113	97
2	Number of members	25	30	20	75
	Average loss per member	113	111	116	113
Total	Number of members				100
	Average loss per member				109

$$\sum_{i=1}^{2}\sum_{t=1}^{3} m_{it}(x_{it}-\bar{x}_i)^2 = 2{,}020$$

$$\sum_{i=1}^{2} m_i(\bar{x}_i-\bar{x})^2 = 4{,}800$$

(a) Determine the nonparametric empirical Bayes premium for Group 1, using $\hat{\mu} = \overline{X}$ as the complement of credibility.

(b) Determine the nonparametric empirical Bayes premium for Group 1, using the method that preserves total losses. [Course 4 − Spring 2001 − #32]

18. You are making credibility estimates for regional rating factors. You observe that the Bühlmann-Straub nonparametric empirical Bayes method can be applied, with rating factor playing the role of pure premium.

X_{it} denotes the rating factor for Region i and Year t, where $i = 1, 2, 3$ and $t = 1, 2, 3, 4$. Corresponding to each rating factor is the number of reported claims, m_{it}, measuring exposure.

i	$m_i = \sum_{t=1}^{4} m_{it}$	$\overline{X}_i = \frac{1}{m_i}\sum_{t=1}^{4} m_{it}X_{it}$	$\hat{v}_i = \frac{1}{3}\sum_{t=1}^{4} m_{it}(X_{it}-\overline{X}_i)^2$	$m_i(\overline{X}_i-\overline{X})^2$
1	50	1.406	0.536	0.887
2	300	1.298	0.125	0.191
3	150	1.178	0.172	1.348

(a) Determine the credibility estimate of the rating factor for Region 1 using $\hat{\mu} = \overline{X}$ as the complement of credibility.

(b) Determine the credibility estimate of the rating factor for Region 1 using the method that preserves $\sum_{i=1}^{3} m_i\overline{X}_i$. [Course 4 – Fall 2001 − #30]

19. You are given total claims for two policyholders:

	Year			
Policyholder	1	2	3	4
X	730	800	650	700
Y	655	650	625	750

Using the nonparametric empirical Bayes Method, determine the Bühlmann credibility premium for Policyholder Y. [Course 4 − Fall 2003 − #15]

2.2 Semiparametric Estimation

20. Two insurance policies produced the following claims during a four-year period:

		Year			
Insured		***Y***	***Y*+1**	***Y*+2**	***Y*+3**
A	Number of Claims	3	1	0	2
	Insured Vehicles	3	2	2	2
B	Number of Claims		0	1	1
	Insured Vehicles		3	3	4

Assume that the number of claims for each vehicle each year has a Poisson distribution and that each vehicle on a policy has the same expected claim frequency.

Estimate the expected annual number of claims per vehicle for each insured using semiparametric empirical Bayes estimation.

21. Assume that the number of claims a driver has during the year is Poisson distributed with an unknown mean that varies by driver. The experience for 100 drivers for one year is as follows:

Number of Claims during the Year	**Number of Drivers**
0	54
1	33
2	10
3	2
4	1

Determine the credibility of one year's experience for a single driver using semiparametric empirical Bayes estimation. [Course 4 − Spring 2000 − #33]

22. The following information comes from a study of robberies of convenience stores over the course of a year:

(i) X_i is the number of robberies of the i^{th} store, with $i = 1, 2, \ldots, 500$.

(ii) $\sum_{i=1}^{500} X_i = 50$

(iii) $\sum_{i=1}^{500} X_i^2 = 220$

(iv) The number of robberies of a given store during the year is assumed to be Poisson distributed with an unknown mean that varies by store.

Determine the semiparametric empirical Bayes estimate of the expected number of robberies next year of a store that reported no robberies during the studied year. [Course 4 − Fall 2000 − #7]

Solutions to Exercises

1. Credibility Models

1. $\mu = E[Z] = E[w_1X + w_2Y] = w_1E[X] + w_2E[Y] = w_1\mu + w_2\mu \rightarrow w_1 + w_2 = 1$

   $Var[Z] = Var[w_1X + w_2Y\,] = w_1^2\sigma_X^2 + w_2^2\sigma_Y^2 = w_1^2\sigma_X^2 + (1-w_1)^2\sigma_Y^2$

   $$\frac{\partial}{\partial w_1}Var[Z] = \frac{\partial}{\partial w_1}\left\{w_1^2\sigma_X^2 + (1-w_1)^2\sigma_Y^2\right\} = 2w_1\sigma_X^2 - 2(1-w_1)\sigma_Y^2 = 0 \;\rightarrow w_1 = \frac{\sigma_Y^2}{\sigma_X^2 + \sigma_Y^2}$$

   Multiply the numerator and denominator of w_1 by $\dfrac{1}{\sigma_X^2\sigma_Y^2}$ to get $w_1 = \dfrac{1/\sigma_X^2}{1/\sigma_X^2 + 1/\sigma_Y^2}$.

   $$w_2 = 1 - w_1 = 1 - \frac{\sigma_Y^2}{\sigma_X^2 + \sigma_Y^2} = \frac{\sigma_X^2}{\sigma_X^2 + \sigma_Y^2} = \frac{1/\sigma_Y^2}{1/\sigma_X^2 + 1/\sigma_Y^2}\ .$$

2. (a) $E_{X\mid\Theta}[X \mid \Theta = \text{A}\,] = .6(0) + .3(2) + .1(4) =$ **1.0** , $E_{X\mid\Theta}[X \mid \Theta = \text{B}\,] = .1(0) + .3(2) + .6(4) =$ **3.0**
   (b) $E[X] = E_\Theta[E_{X\mid\Theta}[X \mid \Theta]] = (1/2)(1.0) + (1/2)(3.0) = 2.0$,
   $Var_\Theta[E_{X\mid\Theta}[X \mid \Theta]] = (1/2)(1.0 - 2.0)^2 + (1/2)(3.0 - 2.0)^2 =$ **1.0**
   (c) $Var_{X\mid\Theta}[X \mid \Theta = \text{A}\,] = .6(0 - 1.0)^2 + .3(2 - 1.0)^2 + .1(4 - 1.0)^2 =$ **1.8**
   $Var_{X\mid\Theta}[X \mid \Theta = \text{B}\,] = .1(0 - 3.0)^2 + .3(2 - 3.0)^2 + .6(4 - 3.0)^2 =$ **1.8**
   (d) $E_\Theta[Var_{X\mid\Theta}[X \mid \Theta]] = (1/2)(1.8) + (1/2)(1.8) =$ **1.8**
   (e) $Var[X] = (1/2)[.6(0 - 2.0)^2 + .3(2 - 2.0)^2 + .1(4 - 2.0)^2]$
   $+ (1/2)\,[.1(0 - 2.0)^2 + .3(2 - 2.0)^2 + .6(4 - 2.0)^2] =$ **2.8 = 1.0 + 1.8**

1.1 Bühlmann Model

3. Variance of the hypothetical means X = amount of claim
   $E[X \mid$ Risk 1] = .5(250) + .3(2,500) + .2(60,000) = 12,875
   $E[X \mid$ Risk 2] = .7(250) + .2(2,500) + .1(60,000) = 6,675
   $E[X]$ = (2/3)(12,875) + (1/3)(6,675) = 10,808.33
   $VHM = (2/3)(12{,}875 - 10{,}808.33)^2 + (1/3)(6{,}675 - 10{,}808.33)^2 = 8{,}542{,}222.2$

   Expected value of the process variance
   $Var[X \mid \text{Risk 1}\,] = .5(250 - 12{,}875)^2 + .3(2{,}500 - 12{,}875)^2 + .2(60{,}000 - 12{,}875)^2 =$
   556,140,625.0
   $Var[X \mid \text{Risk 2}\,] = .7(250 - 6{,}675)^2 + .2(2{,}500 - 6{,}675)^2 + .1(60{,}000 - 6{,}675)^2 =$
   316,738,125.0
   EPV = (2/3)(556,140,625.0) + (1/3)(316,738,125.0) = 476,339,791.7

   $K = EPV / VHM$ = 476,339,791.7 / 8,542,222.2 = 55.76
   $Z = N / (N + K)$ = 1 / (1 + 55.76) = 1 / 56.76

Bühlmann credibility estimate = (1 / 56.76)(250) + (55.76 / 56.76)(10,808.33) = **10,622**

1.2 Bühlmann-Straub Model

4. (a) $E\left[\frac{X_1 + X_2 + ... + X_n}{n}\right] = \left(\frac{1}{n}\right)\sum_{j=1}^{n} E[X_j] = E[X_j] = .6(0) + .3(2) + .1(4)$ = **1.0**

(b) $Var\left[\frac{X_1 + X_2 + ... + X_n}{n}\right] = \left(\frac{1}{n^2}\right)\sum_{j=1}^{n} Var[X_j] = \frac{Var[X_j]}{n} = \left(\frac{1}{n}\right)\left(E[X_j^2] - E[X_j]^2\right) =$

$[.6(0)^2 + .3(2)^2 + .1(4)^2 - (1.0)^2] / n$ = **1.8 / *n***

5.

Urn	Mean	Variance	
		One Ball	Two Balls
A	1.0	1.8	1.8 / 2 = .9
B	3.0	1.8	1.8 / 2 = .9
Average	2.0	1.8	.9

(a) *EPV* = **.9**
(b) $VHM = (1/2)(1.0)^2 + (1/2)(3.0)^2 - (2.0)^2$ = **1.0**

6. The *EPV* for drawing one ball is 1.8 and the *VHM* (not dependent on the number of balls) is 1.0. $\bar{x}$ = [2(2) + 4(1.5)] / 6 = 5/3 , μ = 2.0 , and
Z = 6 / (6 + (1.8/1.0)) = .7692 → Estimate = .7692(5/3) + (1 − .7692)(2.0) = **1.7436**

7. Let θ be the Poisson parameter.
$EPV = E_\Theta[Var$(annual number of claims for one car)$] = E_\Theta[\Theta] = 1/2$
$VHM = Var_\Theta[E$(annual number of claims for one car)$] = Var_\Theta[\Theta] = E_\Theta[\Theta^2] - (E_\Theta[\Theta])^2 = 1/3 - 1/4 = 1/12$
$K = (1/2)/(1/12) = 6$, $m = 4 + 5 + 2 = 11$, and $Z = 11 / (11 + 6) = 11/17$
$\bar{x} = 3/11$, $\mu = 1/2$
Estimated claims per car = (11/17)(3/11)+(6/17)(1/2) = 6 /17
→ Estimated claims for three cars = 3(6/17) = **18/17 = 1.0588**

8.

Class	Probability	Mean	Variance
I	¼	.1	.09
II	¼	.2	.16
III	¼	.5	.25
IV	¼	.9	.09
		μ = .4250	*EPV* = .1475

$VHM = (1/4)[.1^2 + .2^2 + .5^2 + .9^2] - .4250^2 = .096875$

$Z = 4 / [4 + (.1475)/(.096875)] = .7243$, $\bar{x} = 2/4 = 1/2$
Estimate for five insureds = 5{.7243(1/2) + (1−.7243)(.4250)} = **2.3966**

9. $K = 8000/40 = 200$, $\bar{x} = \dfrac{800(15)+600(10)+400(5)}{1800} = \dfrac{100}{9}$, $Z = \dfrac{1800}{1800+200} = .9$

$\mu = 20$ → Estimate = .9(100/9) + .1(20) = **12**

10. (a) $E[\bar{X}] = E[\sum_{i=1}^{n} w_i X_i] = \sum_{i=1}^{n} w_i E[X_i] = \sum_{i=1}^{n} w_i \mu = \left(\sum_{i=1}^{n} w_i\right)\mu = \mu$

(b) $Var[\bar{X}] = Var[\sum_{i=1}^{n} w_i X_i] = \sum_{i=1}^{n} Var[w_i X_i] = \sum_{i=1}^{n} w_i^2 Var[X_i] = \sum_{i=1}^{n} w_i^2 \dfrac{\sigma^2}{m_i} = \sigma^2 \sum_{i=1}^{n} \dfrac{w_i^2}{m_i}$

Minimizing $Var[\bar{X}]$ is equivalent to minimizing $\sum_{i=1}^{n} \dfrac{w_i^2}{m_i}$ with the constraint $\sum_{i=1}^{n} w_i = 1$. Lagrange multipliers are a convenient tool to use here, though not the only way to find the minimum.

$$\nabla\left(\sum_{i=1}^{n}\frac{w_i^2}{m_i} + \lambda\sum_{i=1}^{n} w_i\right) = 0 \quad \rightarrow \quad \frac{\partial}{\partial w_k}\left(\sum_{i=1}^{n}\frac{w_i^2}{m_i} + \lambda\sum_{i=1}^{n} w_i\right) = 0 \quad \rightarrow \quad \frac{2w_k}{m_k} + \lambda = 0$$

$\rightarrow w_k = -\dfrac{\lambda m_k}{2}$. Because $\sum_{i=1}^{n} w_i = 1$, then $\sum_{i=1}^{n}\left(-\dfrac{\lambda m_i}{2}\right) = 1$, or $-\dfrac{\lambda}{2}\sum_{i=1}^{n} m_i = 1$.

Recall that $\sum_{i=1}^{n} m_i = m$ which leads to $\lambda = -2/m$. The result is:

$$w_k = -\frac{\lambda m_k}{2} = -\frac{(-2/m)m_k}{2} = \frac{m_k}{m}$$

11. Variance of the hypothetical means: λ is the mean number of claims for any insured. The prior distribution for λ is gamma:
$E[\lambda] = (1/100)\Gamma(6+1) / \Gamma(6) = 6/100$, $E[\lambda^2] = (1/100)^2\,\Gamma(6+2) / \Gamma(6) = 42 / (100)^2$, and
$Var[\lambda] = E[\lambda^2] - (E[\lambda])^2 = 42 / (100)^2 - (6/100)^2 = 6 / (100)^2 = VHM$.

Expected value of the process variance: Because the distribution is Poisson, the variance of the number of claims equals the mean λ, so λ is the process variance of the number of claims for an insured and $E[\lambda] = (1/100)\Gamma(6+1) / \Gamma(6) = 6/100 = EPV$.

$K = EPV / VHM = (6/100) / (6/100^2) = 100$
$m = 100 + 150 + 200 = 450$
$Z = 450 / (450 + 100) = 9 / 11$
$\bar{x} = (6+8+11)/450 = 1/18$
$\mu = E[\lambda] = 6/100$
Bühlmann-Straub credibility estimate for 300 insureds
$= 300\,[(9/11)(1/18) + (1 - 9/11)(6/100)] =$ **16.91**

2.1 Nonparametric Estimation

12. $N = 4$ and $R = 2$

$$\bar{x}_1 = \frac{1+0+1+0}{4} = \frac{1}{2}, \ \bar{x}_2 = \frac{2+3+3+1}{4} = \frac{9}{4}, \ \bar{x} = \left(\frac{1}{2}+\frac{9}{4}\right)/2 = \frac{11}{8}$$

$$\hat{\sigma}_1^2 = \left(\frac{1}{3}\right)\left[\left(1-\frac{1}{2}\right)^2 + \left(0-\frac{1}{2}\right)^2 + \left(1-\frac{1}{2}\right)^2 + \left(0-\frac{1}{2}\right)^2\right] = \frac{1}{3}$$

$$\hat{\sigma}_2^2 = \left(\frac{1}{3}\right)\left[\left(2-\frac{9}{4}\right)^2 + \left(3-\frac{9}{4}\right)^2 + \left(3-\frac{9}{4}\right)^2 + \left(1-\frac{9}{4}\right)^2\right] = \frac{11}{12}$$

(a) $E\hat{P}V = \left(\frac{1}{2}\right)\left(\frac{1}{3}+\frac{11}{12}\right) = \frac{5}{8}$

(b) $V\hat{H}M = \left(\frac{1}{1}\right)\left[\left(\frac{1}{2}-\frac{11}{8}\right)^2 + \left(\frac{9}{4}-\frac{11}{8}\right)^2\right] - \frac{E\hat{P}V}{4} = \frac{11}{8}$

(c) $\hat{Z} = \dfrac{4}{4+\dfrac{5/8}{11/8}} = \dfrac{44}{49}$, Estimate $1 = \frac{44}{49}\left(\frac{1}{2}\right) + \frac{5}{49}\left(\frac{11}{8}\right) =$ **33/56 = .5893**

Estimate $2 = \frac{44}{49}\left(\frac{9}{4}\right) + \frac{5}{49}\left(\frac{11}{8}\right) =$ **121/56 = 2.1607**

13. $\bar{x}_1 = 4, \bar{x}_2 = 6, \bar{x} = 5, \quad \hat{\sigma}_1^2 = 1, \hat{\sigma}_2^2 = 1$

(a) $E\hat{P}V = \left(\frac{1}{2}\right)(1+1) = 1$

(b) $V\hat{H}M = \left(\frac{1}{1}\right)\left[(4-5)^2 + (6-5)^2\right] - \frac{E\hat{P}V}{3} = \frac{5}{3}$

(c) $\hat{K} = \frac{1}{5/3} = \frac{3}{5}$, $\hat{Z} = \dfrac{3}{3+\frac{3}{5}} = \dfrac{5}{6}$, Estimate $1 = \frac{5}{6}(4) + \frac{1}{6}(5) =$ **25/6 = 4.1667**

$$\text{Estimate } 2 = \frac{5}{6}(6) + \frac{1}{6}(5) = \mathbf{35/6 = 5.8333}$$

14. $E\hat{P}V = \dfrac{33.60}{4(7-1)} = 1.4$, $\quad V\hat{H}M = \dfrac{3.30}{(4-1)} - \dfrac{1.4}{7} = .9 \;\rightarrow\; \hat{Z} = \dfrac{7}{7 + \dfrac{1.4}{.9}}$ = **9/11 = .8182**

15. $m_A = 2 + 2 + 2 + 1 = 7$, $\quad m_B = 4 + 4 + 4 = 12$, $\quad m = 7 + 12 = 19$

$$\bar{x}_A = \frac{3+2+3+1}{7} = \frac{9}{7}, \quad \bar{x}_B = \frac{0+1+1}{12} = \frac{1}{6}, \quad \bar{x} = \frac{9+2}{7+12} = \frac{11}{19}$$

$$\hat{\sigma}_A^2 = \frac{2\left(\frac{3}{2} - \frac{9}{7}\right)^2 + 2\left(\frac{2}{2} - \frac{9}{7}\right)^2 + 2\left(\frac{3}{2} - \frac{9}{7}\right)^2 + 1\left(\frac{1}{1} - \frac{9}{7}\right)^2}{4-1} = .1429$$

$$\hat{\sigma}_B^2 = \frac{4\left(\frac{0}{4} - \frac{1}{6}\right)^2 + 4\left(\frac{1}{4} - \frac{1}{6}\right)^2 + 4\left(\frac{1}{4} - \frac{1}{6}\right)^2}{3-1} = .0833$$

Note that one does <u>not</u> have to calculate the separate process variances as above. The two numerators above can be calculated and added and then the total divided by (7 − 2).

$$E\hat{P}V = \frac{(4-1)(.1429) + (3-1)(.0833)}{(4-1) + (3-1)} = .1191$$

$$V\hat{H}M = \frac{7[(9/7) - (11/19)]^2 + 12[(1/6) - (11/19)]^2 - (2-1)(.1191)}{19 - \left(\frac{1}{19}\right)(7^2 + 12^2)} = .6127$$

$$\hat{K} = \frac{.1191}{.6127} = .1944$$

$$\hat{Z}_A = \frac{7}{7 + .1944} = .9730 \rightarrow \quad \hat{\mu}_A = .9730(9/7) + (1 - .9730)(11/19) = \mathbf{1.2666}$$

$$\hat{Z}_B = \frac{12}{12 + .1944} = .9841 \rightarrow \quad \hat{\mu}_B = .9841(1/6) + (1 - .9841)(11/19) = \mathbf{.1732}$$

Comment: Note that the indications do not balance to the overall mean: {7(1.2666)+12(.1732)}/19 $\neq$ 11/19 though it is close. As explained in the reading, these can be brought into balance if the credibilities are used as weights to compute a weighted overall mean to use for the complement of credibility.

16. $E\hat{P}V = [\ 2{,}000(0-3)^2 + 1{,}000(5-3)^2 + 1{,}000(6-3)^2 + 1{,}000(4-3)^2 + 450(15-10)^2 +$

$$250(2-10)^2 + 175(15-10)^2 + 125(1-10)^2 \ ]/[4+4-2] \ = \ 12{,}291.67$$

(a) $\hat{Z}_{\mathrm{Y}} = \dfrac{1{,}000}{1{,}000 + \dfrac{12{,}291.67}{17.125}} = .5822$, $\quad \bar{x} = \dfrac{5{,}000(3) + 1{,}000(10)}{5{,}000 + 1{,}000} = 4.1667$

Youth Class estimate = .5822(10) + .4178(4.1667) = **7.5628**

(b) $\hat{Z}_{\mathrm{A}} = \dfrac{5{,}000}{5{,}000 + \dfrac{12{,}291.67}{17.125}} = .8745$, $\quad \hat{\mu} = \dfrac{.8745(3) + .5822(10)}{.8745 + .5822} = 5.7977$

Youth Class estimate = .5822(10) + .4178(5.7977) = **8.2443**

17. $E\hat{P}V = \dfrac{2{,}020}{6-2} = 505$

$$V\hat{H}M = \left(4{,}800 - (2-1)505\right) / \left(100 - \left(\frac{1}{100}\right)(25^2 + 75^2)\right) = 114.533$$

$$\hat{K} = \frac{505}{114.533} = 4.4092, \quad \hat{Z}_1 = \frac{25}{25 + 4.4092} = .8501$$

(a) Estimate for Group 1 = .8501(97) + .1499(109) = **98.7988**

(b) $\hat{Z}_2 = \dfrac{75}{75 + 4.4092} = .9445$, $\quad \hat{\mu} = \dfrac{.8501(97) + .9445(113)}{.8501 + .9445} = 105.4208$

Estimate for Group 1 = .8501(97) + .1499(105.4208) = **98.2623**

18. $E\hat{P}V = \dfrac{3(.536) + 3(.125) + 3(.172)}{3+3+3} = .2777$

$$V\hat{H}M = \frac{(.887 + .191 + 1.348) - (3-1).2777}{500 - \left(\frac{1}{500}\right)\left(50^2 + 300^2 + 150^2\right)} = .006928, \quad \hat{K} = \frac{.2777}{.006928} = 40.08$$

$$\hat{Z}_1 = \frac{50}{50 + 40.08} = .5551, \quad \bar{x} = \frac{50(1.406) + 300(1.298) + 150(1.178)}{50 + 300 + 150} = 1.273$$

(a) Estimate for Region 1 = .5551(1.406) + .4449(1.273) = **1.347**

(b) $\hat{Z}_2 = \dfrac{300}{300 + 40.08} = .8821$, $\hat{Z}_3 = \dfrac{150}{150 + 40.08} = .7891$

$$\hat{\mu} = \frac{.5551(1.406) + .8821(1.298) + .7891(1.178)}{.5551 + .8821 + .7891} = 1.2824$$

Estimate for Region 1 = .5551(1.406) + .4449(1.2824) = **1.351**

19. $\bar{x} = (730 + 800 + 650 + 700)/4 = 720$, $\bar{y} = (655 + 650 + 625 + 750)/4 = 670$, and $\hat{\mu} = (720 + 670)/2 = 695$.

$$\hat{\sigma}_{\mathrm{X}}^2 = \frac{(730-720)^2 + (800-720)^2 + (650-720)^2 + (700-720)^2}{4-1} = 11{,}800/3$$

$$\hat{\sigma}_{\mathrm{Y}}^2 = \frac{(655-670)^2 + (650-670)^2 + (625-670)^2 + (750-670)^2}{4-1} = 9{,}050/3$$

$$E\hat{P}V = [(11{,}800/3) + (9{,}050/3)]/2 = 3{,}475$$

$$V\hat{H}M = [(720-695)^2 + (670-695)^2]/(2-1) - (3{,}475/4) = 381.25$$

$$\hat{Z} = \frac{4}{4 + [3{,}475/381.25]} = .3050$$

Bühlmann credibility estimate for Y = .3050(670) + (1−.3050)(695) = **687.375**

2.2 Semiparametric Estimation

20. $\bar{x}_{\mathrm{A}} = \frac{3+1+0+2}{3+2+2+2} = \frac{2}{3}$, $\bar{x}_{\mathrm{B}} = \frac{0+1+1}{3+3+4} = \frac{1}{5}$, $\bar{x} = \frac{6+2}{19} = \frac{8}{19}$

With the Poisson assumption: $E\hat{P}V = \bar{x} = \frac{8}{19}$

$$V\hat{H}M = \frac{9\left(\frac{2}{3} - \frac{8}{19}\right)^2 + 10\left(\frac{1}{5} - \frac{8}{19}\right)^2 - (2-1)\left(\frac{8}{19}\right)}{19 - \left(\frac{1}{19}\right)\left(9^2 + 10^2\right)} = .06444,\quad \hat{K} = \frac{8/19}{.06444} = 6.5340$$

$$\hat{Z}_{\mathrm{A}} = \frac{9}{9 + 6.5340} = .5794,\qquad \hat{Z}_{\mathrm{B}} = \frac{10}{10 + 6.5340} = .6048$$

Estimate for A = .5794(2/3) + .4206(8/19) = **.5634**

Estimate for B = .6048(1/5) + .3952(8/19) = **.2874**

Comment: Of course, answers that balance to the total can be derived by calculating a complement of credibility using the credibility weights.

21. Note R = 100 and N = 1. Since the claims process is Poisson:

$$E\hat{P}V = \bar{x} = \frac{54(0)+33(1)+10(2)+2(3)+1(4)}{100} = .63$$

$$V\hat{H}M = \frac{54(0-.63)^2+33(1-.63)^2+10(2-.63)^2+2(3-.63)^2+1(4-.63)^2}{99} - \frac{.63}{1} = .05$$

$$\hat{Z} = \frac{1}{1+\frac{.63}{.05}} = \mathbf{.0735}$$

22. The identity $\sum_{i=1}^{n}(X_i - \bar{X})^2 = \sum_{i=1}^{n} X_i^2 - n\bar{X}^2$ should be used in the calculation of the $V\hat{H}M$.

Note that R = 500 and N = 1: $\bar{X} = \frac{\sum_{i=1}^{500} X_i}{500}$, $\bar{x} = \frac{50}{500} = .10$. Since the distribution is Poisson,

$$E\hat{P}V = \bar{x} = .10$$

$$V\hat{H}M = \left(\frac{1}{499}\right)(220 - 500(.10)^2) - .10 = .3309, \quad \hat{Z} = \frac{1}{1+\frac{.10}{.3309}} = .7679$$

Estimate = .7679(0) + .2321(.10) = **.0232**