CHAPTER 8

CREDIBILITY

HOWARD C. MAHLER AND CURTIS GARY DEAN

1. INTRODUCTION

Credibility theory provides tools to deal with the randomness of data that is used for predicting future events or costs. For example, an insurance company uses past loss information of an insured or group of insureds to estimate the cost to provide future insurance coverage. But, insurance losses arise from random occurrences. The average annual cost of paying insurance losses in the past few years may be a poor estimate of next year's costs. The expected accuracy of this estimate is a function of the variability in the losses. This data by itself may not be acceptable for calculating insurance rates.

Rather than relying solely on recent observations, better estimates may be obtained by combining this data with other information. For example, suppose that recent experience indicates that Carpenters should be charged a rate of \$5 (per \$100 of payroll) for workers compensation insurance. Assume that the current rate is \$10. What should the new rate be? Should it be \$5, \$10, or somewhere in between? Credibility is used to weight together these two estimates.

The basic formula for calculating credibility weighted estimates is:

$$\text{Estimate} = Z \times [\text{Observation}] + (1 - Z) \times [\text{Other Information}],$$
$$0 \leq Z \leq 1.$$

Z is called the credibility assigned to the observation. $1 - Z$ is generally referred to as the complement of credibility. If the body of observed data is large and not likely to vary much from one period to another, then Z will be closer to one. On the other hand, if the observation consists of limited data, then Z will be closer to zero and more weight will be given to other information.

The current rate of \$10 in the above example is the "Other Information." It represents an estimate or prior hypothesis of a rate to charge in the absence of the recent experience. As recent experience becomes available, then an updated estimate combining the recent experience and the prior hypothesis can be calculated. Thus, the use of credibility involves a linear estimate of the true expectation derived as a result of a compromise between observation and prior hypothesis. The Carpenters' rate for workers compensation insurance is $Z \times \$5 + (1 - Z) \times \$10$ under this model.

Following is another example demonstrating how credibility can help produce better estimates:

Example 1.1: In a large population of automobile drivers, the average driver has one accident every five years or, equivalently, an annual frequency of .20 accidents per year. A driver selected randomly from the population had three accidents during the last five years for a frequency of .60 accidents per year. What is your estimate of the expected future frequency rate for this driver? Is it .20, .60, or something in between?

[*Solution:* If we had no information about the driver other than that he came from the population, we should go with the .20. However, we know that the driver's observed frequency was .60. Should this be our estimate for his future accident frequency? Probably not. There is a correlation between prior accident frequency and future accident frequency, but they are not perfectly correlated. Accidents occur randomly and even good drivers with low expected accident frequencies will have accidents. On the other hand, bad drivers can go several years without an accident. A better answer than either .20 or .60 is most likely something in between: this driver's Expected Future Accident Frequency $= Z \times .60 + (1 - Z) \times .20$.]

The key to finishing the solution for this example is the calculation of Z. How much credibility should be assigned to the information known about the driver? The next two sections explain the calculation of Z.

First, the classical credibility model will be covered in Section 2. It is also referred to as limited fluctuation credibility because it attempts to limit the effect that random fluctuations in the observations will have on the estimates. The credibility Z is a function of the expected variance of the observations versus the selected variance to be allowed in the first term of the credibility formula, $Z \times$ [Observation].

Next Bühlmann credibility is described in Section 3. This model is also referred to as least squares credibility. The goal with this approach is the minimization of the square of the error between the estimate and the true expected value of the quantity being estimated.

Credibility theory depends upon having prior or collateral information that can be weighted with current observations. Another approach to combining current observations with prior information to produce a better estimate is Bayesian analysis. Bayes Theorem is the foundation for this analysis. This is covered is Section 4. It turns out that Bühlmann credibility estimates are the best linear least squares fits to Bayesian estimates. For this reason Bühlmann credibility is also referred as Bayesian credibility.

In some situations the resulting formulas of a Bayesian analysis exactly match those of Bühlmann credibility estimation; that is, the Bayesian estimate is a linear weighting of current and prior information with weights Z and $(1-Z)$ where Z is the Bühlmann credibility. In Section 5 this is demonstrated in the important special case of the Gamma-Poisson frequency process.

The last section discusses practical issues in the application of credibility theory including some examples of how to calculate credibility parameters.

The Appendices include basic facts on several frequency and severity distributions and the solutions to the exercises.

2. CLASSICAL CREDIBILITY

2.1. Introduction

In Classical Credibility, one determines how much data one needs before one will assign to it 100% credibility. This amount of data is referred to as the **Full Credibility Criterion or the Standard for Full Credibility**. If one has this much data or more, then $Z = 1.00$; if one has observed less than this amount of data then $0 \leq Z < 1$.

For example, if we observed 1,000 full-time Carpenters, then we might assign 100% credibility to their data.[1] Then if we observed 2,000 full-time Carpenters we would also assign them 100% credibility. 100 full-time Carpenters might be assigned 32% credibility. In this case the observation has been assigned **partial credibility**, i.e., less than full credibility. Exactly how to determine the amount of credibility assigned to different amounts of data is discussed in the following sections.

There are four basic concepts from Classical Credibility which will be covered:

1. How to determine the criterion for *Full Credibility* when estimating *frequencies*;
2. How to determine the criterion for *Full Credibility* when estimating *severities*;
3. How to determine the criterion for *Full Credibility* when estimating *pure premiums* (loss costs);
4. How to determine the amount of *partial credibility* to assign when one has less data than is needed for full credibility.

[1] For workers compensation that data would be dollars of loss and dollars of payroll.

Example 2.1.1: The observed claim frequency is 120. The credibility given to this data is 25%. The complement of credibility is given to the prior estimate of 200. What is the new estimate of the claim frequency?

[*Solution:* $.25 \times 120 + (1 - .25) \times 200 = 180$.]

2.2. Full Credibility for Frequency

Assume we have a Poisson process for claim frequency, with an average of 500 claims per year. Then, the observed numbers of claims will vary from year to year around the mean of 500. The variance of a Poisson process is equal to its mean, in this case 500. This Poisson process can be approximated by a Normal Distribution with a mean of 500 and a variance of 500.

The Normal Approximation can be used to estimate how often the observed results will be far from the mean. For example, how often can one expect to observe more than 550 claims? The standard deviation is $\sqrt{500} = 22.36$. So 550 claims corresponds to about $50/22.36 = 2.24$ standard deviations greater than average. Since $\Phi(2.24) = .9875$, there is approximately a 1.25% chance of observing more than 550 claims.[2]

Thus there is about a 1.25% chance that the observed number of claims will exceed the expected number of claims by 10% or more. Similarly, the chance of observing fewer than 450 claims is approximately 1.25%. So the chance of observing a number of claims that is outside the range from -10% below to $+10\%$ above the mean number of claims is about 2.5%. In other words, the chance of observing within $\pm 10\%$ of the expected number of claims is 97.5% in this case.

[2]More precisely, the probability should be calculated including the continuity correction. The probability of more than 550 claims is approximately $1 - \Phi((550.5 - 500)/\sqrt{500}) = 1 - \Phi(2.258) = 1 - .9880 = 1.20\%$.

More generally, one can write this algebraically. The probability P that observation X is within $\pm k$ of the mean μ is:

$$\begin{aligned} P &= \mathrm{Prob}[\mu - k\mu \leq X \leq \mu + k\mu] \\ &= \mathrm{Prob}[-k(\mu/\sigma) \leq (X-\mu)/\sigma \leq k(\mu/\sigma)] \end{aligned} \qquad (2.2.1)$$

The last expression is derived by subtracting through by μ and then dividing through by standard deviation σ. Assuming the Normal Approximation, the quantity $u = (X-\mu)/\sigma$ is normally distributed. For a Poisson distribution with expected number of claims n, then $\mu = n$ and $\sigma = \sqrt{n}$. The probability that the observed number of claims N is within $\pm k\%$ of the expected number $\mu = n$ is:

$$P = \mathrm{Prob}[-k\sqrt{n} \leq u \leq k\sqrt{n}]$$

In terms of the cumulative distribution for the unit normal, $\Phi(u)$:

$$\begin{aligned} P &= \Phi(k\sqrt{n}) - \Phi(-k\sqrt{n}) = \Phi(k\sqrt{n}) - (1 - \Phi(k\sqrt{n})) \\ &= 2\Phi(k\sqrt{n}) - 1 \end{aligned}$$

Thus, for the Normal Approximation to the Poisson:

$$P = 2\Phi(k\sqrt{n}) - 1 \qquad (2.2.2)$$

Or, equivalently:

$$\Phi(k\sqrt{n}) = (1+P)/2. \qquad (2.2.3)$$

Example 2.2.1: If the number of claims has a Poisson distribution, compute the probability of being within $\pm 5\%$ of a mean of 100 claims using the Normal Approximation to the Poisson.

[*Solution:* $2\Phi(.05\sqrt{100}) - 1 = 38.29\%$.]

Here is a table showing P, for $k = 10\%$, 5%, 2.5%, 1%, and 0.5%, and for 10, 50, 100, 500, 1,000, 5,000, and 10,000 claims:

Probability of Being Within $\pm k$ of the Mean

Expected # of Claims	$k = 10\%$	$k = 5\%$	$k = 2.5\%$	$k = 1\%$	$k = 0.5\%$
10	24.82%	12.56%	6.30%	2.52%	1.26%
50	52.05%	27.63%	14.03%	5.64%	2.82%
100	68.27%	38.29%	19.74%	7.97%	3.99%
500	97.47%	73.64%	42.39%	17.69%	8.90%
1,000	99.84%	88.62%	57.08%	24.82%	12.56%
5,000	100.00%	99.96%	92.29%	52.05%	27.63%
10,000	100.00%	100.00%	98.76%	68.27%	38.29%

Turning things around, given values of P and k, then one can compute the number of expected claims n_0 such that the chance of being within $\pm k$ of the mean is P. n_0 can be calculated from the formula $\Phi(k\sqrt{n_0}) = (1+P)/2$. Let y be such that $\Phi(y) = (1+P)/2$. Then given P, y is determined from a normal table. Solving for n_0 in the relationship $k\sqrt{n_0} = y$ yields $n_0 = (y/k)^2$. If the goal is to be within $\pm k$ of the mean frequency with a probability at least P, then the Standard for Full Credibility is

$$n_0 = y^2/k^2, \tag{2.2.4}$$

where y is such that

$$\Phi(y) = (1+P)/2. \tag{2.2.5}$$

Here are values of y taken from a normal table corresponding to selected values of P:

P	$(1+P)/2$	y
80.00%	90.00%	1.282
90.00%	95.00%	1.645
95.00%	97.50%	1.960
99.00%	99.50%	2.576
99.90%	99.95%	3.291
99.99%	99.995%	3.891

Example 2.2.2: For $P = 95\%$ and for $k = 5\%$, what is the number of claims required for Full Credibility for estimating the frequency?

[*Solution:* $y = 1.960$ since $\Phi(1.960) = (1 + P)/2 = 97.5\%$. Therefore $n_0 = y^2/k^2 = (1.96/.05)^2 = 1537$.]

Here is a table[3] of values for the Standard for Full Credibility for the frequency n_0, given various values of P and k:

Standards for Full Credibility for Frequency (Claims)

Probability Level P	$k = 30\%$	$k = 20\%$	$k = 10\%$	$k = 7.5\%$	$k = 5\%$	$k = 2.5\%$	$k = 1\%$
80.00%	18	41	164	292	657	2,628	16,424
90.00%	30	68	271	481	**1,082**	4,329	27,055
95.00%	43	96	384	**683**	1,537	6,146	38,415
96.00%	47	105	422	750	1,687	6,749	42,179
97.00%	52	118	471	837	1,884	7,535	47,093
98.00%	60	135	541	962	2,165	8,659	54,119
99.00%	74	166	664	1,180	2,654	10,616	66,349
99.90%	120	271	1,083	1,925	4,331	17,324	108,276
99.99%	168	378	1,514	2,691	6,055	24,219	151,367

The value 1,082 claims corresponding to $P = 90\%$ and $k = 5\%$ is commonly used in applications. For $P = 90\%$ we want to have a 90% chance of being within $\pm k$ of the mean, so we are willing to have a 5% probability outside on either tail, for a total of 10% probability of being outside the acceptable range. Thus $\Phi(y) = .95$ or $y = 1.645$. Thus $n_0 = y^2/k^2 = (1.645/.05)^2 = 1{,}082$ claims.

In practical applications appropriate values of P and k have to be selected.[4] While there is clearly judgment involved in the

[3]See the Table in Longley-Cook's "An Introduction to Credibility Theory" (1962) or "Some Notes on Credibility" by Perryman, *PCAS*, 1932. Tables of Full Credibility standards have been available and used by actuaries for many years.

[4]For situations that come up repeatedly, the choice of P and k may have been made several decades ago, but nevertheless the choice was made at some point in time.

choice of P and k, the Standards for Full Credibility for a given application are generally chosen within a similar range. This same type of judgment is involved in the choice of error bars around a statistical estimate of a quantity. Often ± 2 standard deviations (corresponding to about a 95% confidence interval) will be chosen, but that is not necessarily better than choosing ± 1.5 or ± 2.5 standard deviations. So while Classical Credibility involves somewhat arbitrary judgments, that has not stood in the way of its being very useful for decades in many applications.

Subsequent sections deal with estimating severities or pure premiums rather than frequencies. As will be seen, in order to calculate a Standard for Full Credibility for severities or the pure premium, generally one first calculates a Standard for Full Credibility for the frequency.

Variations from the Poisson Assumptions

If one desires that the chance of being within $\pm k$ of the mean frequency to be at least P, then the Standard for Full Credibility is $n_0 = y^2/k^2$, where y is such that $\Phi(y) = (1 + P)/2$.

However, this depended on the following assumptions:

1. One is trying to estimate frequency;
2. Frequency is given by a Poisson process (so that the variance is equal to the mean);
3. There are enough expected claims to use the Normal Approximation to the Poisson process.

Occasionally, a Binomial or Negative Binomial Distribution will be substituted for a Poisson distribution, in which case the difference in the derivation is that the variance is not equal to the mean.

For example, assume one has a Binomial Distribution with parameters $n = 1{,}000$ and $p = .3$. The mean is 300 and the variance is (1,000)(.3)(.7) = 210. So the chance of being within ±5% of the expected value is approximately:

$$\Phi((.05)(300)/210^{.5}) - \Phi((-.05)(300)/210^{.5})$$
$$\simeq \Phi(1.035) - \Phi(-1.035) \simeq .8496 - .1504 \simeq 69.9\%.$$

So, in the case of a Binomial with parameter .3, the Standard for Full Credibility with $P = 70\%$ and $k = \pm 5\%$ is about 1,000 exposures or 300 expected claims.

If instead a Negative Binomial Distribution had been assumed, then the variance would have been greater than the mean. This would have resulted in a standard for Full Credibility greater than in the Poisson situation.

One can derive a more general formula when the Poisson assumption does not apply. The Standard for Full Credibility for Frequency is:[5]

$$\{y^2/k^2\}(\sigma_f^2/\mu_f) \qquad (2.2.6)$$

There is an "extra" factor of the variance of the frequency divided by its mean. This reduces to the Poisson case when $\sigma_f^2/\mu_f = 1$.

Exposures vs. Claims

Standards for Full Credibility are calculated in terms of the expected number of claims. It is common to translate these into a number of exposures by dividing by the (approximate) expected claim frequency. So for example, if the Standard for Full Credibility is 1,082 claims ($P = 90\%$, $k = 5\%$) and the expected claim frequency in Homeowners Insurance were .04 claims per house-year, then $1{,}082/.04 \simeq 27{,}000$ house-years would be a corresponding Standard for Full Credibility in terms of exposures.

Example 2.2.3: E represents the number of homogeneous exposures in an insurance portfolio. The claim frequency *rate* per exposure is a random variable with mean = 0.025 and variance =

[5] A derivation of this formula can be found in Mayerson, et al. "The Credibility of the Pure Premium."

0.0025. A full credibility standard is devised that requires the observed sample frequency *rate* per exposure to be within 5% of the expected population frequency rate per exposure 90% of the time. Determine the value of E needed to produce full credibility for the portfolio's experience.

[*Solution:* First calculate the number of claims for full credibility when the mean does not equal the variance of the frequency: $\{1.645^2/(.05)^2\}\{.0025/.025\} = 108.241$. Then, convert this into exposures by dividing by the claim frequency rate per exposure: $108.241/.025 = 4{,}330$ exposures.]

2.2. *Exercises*

2.2.1. How many claims are required for Full Credibility if one requires that there be a 99% chance of the estimated frequency being within $\pm 2.5\%$ of the true value?

2.2.2. How many claims are required for Full Credibility if one requires that there be a 98% chance of the estimated frequency being within $\pm 7.5\%$ of the true value?

2.2.3. The full credibility standard for a company is set so that the total number of claims is to be within 6% of the true value with probability P. This full credibility standard is calculated to be 900 claims. What is the value of P?

2.2.4. Y represents the number of independent homogeneous *exposures* in an insurance portfolio. The claim frequency *rate* per exposure is a random variable with mean = 0.05 and variance = 0.09. A full credibility standard is devised that requires the observed sample frequency *rate* per exposure to be within 2% of the expected population frequency rate per exposure 94% of the time. Determine the value of Y needed to produce full credibility for the portfolio's experience.

2.2.5. Assume you are conducting a poll relating to a single question and that each respondent will answer either yes or no. You pick a random sample of respondents out of a very large population. Assume that the true percentage of yes responses in the total population is between 20% and 80%. How many respondents do you need, in order to require that there be a 95% chance that the results of the poll are within $\pm 7\%$ of the true answer?

2.2.6. A Standard for Full Credibility has been established for frequency assuming that the frequency is Poisson. If instead the frequency is assumed to follow a Negative Binomial with parameters $k = 12$ and $p = .7$, what is the ratio of the revised Standard for Full Credibility to the original one? (For a Negative Binomial, mean $= k(1-p)/p$ and variance $= k(1-p)/p^2$.)

2.2.7. Let X be the number of claims needed for full credibility, if the estimate is to be within 5% of the true value with a 90% probability. Let Y be the similar number using 10% rather than 5%. What is the ratio of X divided by Y?

2.3. Full Credibility for Severity

The Classical Credibility ideas also can be applied to estimating claim severity, the average size of a claim.

Suppose a sample of N claims, $X_1, X_2, \ldots X_N$, are each independently drawn from a loss distribution with mean μ_s and variance σ_s^2. The severity, i.e. the mean of the distribution, can be estimated by $(X_1 + X_2 + \cdots + X_N)/N$. The variance of the observed severity is $\mathrm{Var}(\sum X_i/N) = (1/N^2)\sum \mathrm{Var}(X_i) = \sigma_s^2/N$. Therefore, the standard deviation for the observed severity is $\sigma_s/\sqrt{N}$.

The probability that the observed severity S is within $\pm k$ of the mean μ_s is:

$$P = \mathrm{Prob}[\mu_s - k\mu_s \leq S \leq \mu_s + k\mu_s]$$

Subtracting through by the mean μ_s, dividing by the standard deviation $\sigma_s/\sqrt{N}$, and substituting u in for $(S-\mu_s)/(\sigma_s/\sqrt{N})$ yields:

$$P = \mathrm{Prob}[-k\sqrt{N}(\mu_s/\sigma_s) \leq u \leq k\sqrt{N}(\mu_s/\sigma_s)]$$

This is identical to the frequency formula in Section 2.2 except for the additional factor of (μ_s/σ_s).

According to the Central Limit Theorem, the distribution of observed severity $(X_1 + X_2 + \cdots + X_N)/N$ can be approximated by a normal distribution for large N. Assume that the Normal Approximation applies and, as before with frequency, define y such that $\Phi(y) = (1 + P)/2$. In order to have probability P that the observed severity will differ from the true severity by less than $\pm k\mu_s$, we want $y = k\sqrt{N}(\mu_s/\sigma_s)$. Solving for N:

$$N = (y/k)^2(\sigma_s/\mu_s)^2 \qquad (2.3.1)$$

The ratio of the standard deviation to the mean, $(\sigma_s/\mu_s) = CV_S$, is the coefficient of variation of the claim size distribution. Letting n_0 be the full credibility standard for frequency given P and k produces:

$$N = n_0 CV_S^2 \qquad (2.3.2)$$

This is the ***Standard for Full Credibility for Severity.***

Example 2.3.1: The coefficient of variation of the severity is 3. For $P = 95\%$ and $k = 5\%$, what is the number of claims required for Full Credibility for estimating the severity?

[*Solution:* From Example 2.2.2, $n_0 = 1537$. Therefore, $N = 1537(3)^2 = 13,833$ claims.]

2.3. Exercises

2.3.1. The claim amount distribution has mean 1,000 and variance 6,000,000. Find the number of claims required for full credibility if you require that there will be a 90% chance that the estimate of severity is correct within $\pm 1\%$.

2.3.2. The Standard for Full Credibility for Severity for claim distribution A is N claims for a given P and k. Claim distribution B has the same mean as distribution A, but a standard deviation that is twice as large as A's. Given the same P and k, what is the Standard for Full Credibility for Severity for distribution B?

2.4. Process Variance of Aggregate Losses, Pure Premiums, and Loss Ratios

Suppose that N claims of sizes $X_1, X_2, \ldots, X_N$ occur during the observation period. The following quantities are useful in analyzing the cost of insuring a risk or group of risks:

$$\text{Aggregate Losses}: \quad L = (X_1 + X_2 + \cdots + X_N)$$

$$\text{Pure Premium}: \quad PP = (X_1 + X_2 + \cdots + X_N)/\text{Exposures}$$

$$\text{Loss Ratio}: \quad LR = (X_1 + X_2 + \cdots + X_N)/\text{Earned Premium}$$

We'll work with the Pure Premium in this section, but the development applies to the other two as well.

Pure Premiums are defined as losses divided by exposures.[6] For example, if 200 cars generate claims that total to \$80,000 during a year, then the observed Pure Premium is \$80,000/200 or \$400 per car-year. Pure premiums are the product of frequency and severity. **Pure Premiums = Losses/Exposures = (Number of Claims/Exposures) (Losses/Number of Claims) = (Frequency)(Severity).** Since they depend on both the number of claims and the size of claims, pure premiums have more reasons to vary than do either frequency or severity individually.

Random fluctuation occurs when one rolls dice, spins spinners, picks balls from urns, etc. The observed result varies from time period to time period due to random chance. This is also true for the pure premium observed for a collection of insureds or for an individual insured.[7] The variance of the observed pure premiums for a given risk that occurs due to random fluctuation

[6] The definition of exposures varies by line of insurance. Examples include car-years, house-years, sales, payrolls, etc.

[7] In fact this is the fundamental reason for the existence of insurance.

is referred to as the **process variance**. That is what will be discussed here.[8]

Example 2.4.1: [Frequency and Severity are *not* independent]

Assume the following:

- For a given risk, the number of claims for a single exposure period will be either 0, 1, or 2

Number of Claims	Probability
0	60%
1	30%
2	10%

- If only one claim is incurred, the size of the claim will be 50 with probability 80% or 100 with probability 20%
- If two claims are incurred, the size of each claim, independent of the other, will be 50 with probability 50% or 100 with probability 50%

What is the variance of the pure premium for this risk?

[*Solution:* First list the possible pure premiums and probability of each of the possible outcomes. If there is no claim (60% chance) then the pure premium is zero. If there is one claim, then the pure premium is either 50 with (30%)(80%) = 24% chance or 100 with (30%)(20%) = 6% chance. If there are two claims then there are three possibilities.

Next, the first and second moments can be calculated by listing the pure premiums for all the possible outcomes and taking the weighted average using the probabilities as weights of either the pure premium or its square.

[8]The process variance is distinguished from the variance of the hypothetical pure premiums as discussed in Bühlmann Credibility.

Situation	Probability	Pure Premium	Square of P.P.
0 claims	60.0%	0	0
1 claim @ 50	24.0%	50	2,500
1 claim @ 100	6.0%	100	10,000
2 claims @ 50 each	2.5%	100	10,000
2 claims: 1 @ 50 & 1 @ 100	5.0%	150	22,500
2 claims @ 100 each	2.5%	200	40,000
Overall	100.0%	33	3,575

The average Pure Premium is 33. The second moment of the Pure Premium is 3,575. Therefore, the variance of the pure premium is: $3{,}575 - 33^2 = 2{,}486$.]

Note that the frequency and severity are *not* independent in this example. Rather the severity distribution depends on the number of claims. For example, the average severity is 60 if there is one claim, while the average severity is 75 if there are two claims.

Here is a similar example with independent frequency and severity.

Example 2.4.2: [Frequency and Severity are independent]

Assume the following:

- For a given risk, the number of claims for a single exposure period is given by a binomial distribution with $p = .3$ and $n = 2$.
- The size of a claim will be 50, with probability 80%, or 100, with probability 20%.
- Frequency and severity are independent.

Determine the variance of the pure premium for this risk.

[*Solution:* List the possibilities and compute the first two moments:

Situation	Probability	Pure Premium	Square of P.P.
0	49.00%	0	
1 claim @ 50	33.60%	50	2,500
1 claim @ 100	8.40%	100	10,000
2 claims @ 50 each	5.76%	100	10,000
2 claims: 1 @ 50 & 1 @ 100	2.88%	150	22,500
2 claims @ 100 each	0.36%	200	40,000
Overall	100.0%	36	3,048

Therefore, the variance of the pure premium is: $3{,}048 - 36^2 = 1{,}752$.]

In this second example, since frequency and severity are independent one can make use of the following formula:

Process Variance of Pure Premium =

(Mean Freq.)(Variance of Severity)

+ (Mean Severity)^2(Variance of Freq.)

$$\sigma_{PP}^2 = \mu_f \sigma_S^2 + \mu_S^2 \sigma_f^2 \tag{2.4.1}$$

Note that each of the two terms has a mean and a variance, one from frequency and one from severity. Each term is in dollars squared; that is one way to remember that the mean severity (which is in dollars) enters as a square while that for mean frequency (which is not in dollars) does not.

Example 2.4.3: Calculate the variance of the pure premium for the risk described in Example 2.4.2 using formula (2.4.1).

[*Solution:* The mean frequency is $np = .6$ and the variance of the frequency is $npq = (2)(.3)(.7) = .42$. The average severity is 60 and the variance of the severity is $(.8)(10^2) + (.2)(40^2) =$

400. Therefore the process variance of the pure premium is $(.6)(400) + (60^2)(.42) = 1{,}752$.]

Formula (2.4.1) can also be used to compute the process variance of the aggregate losses and the loss ratio when frequency and severity are independent.

Derivation of Formula (2.4.1)

The above formula for the process variance of the pure premium is a special case of the formula that also underlies analysis of variance:[9]

$$\mathrm{Var}(Y) = E_X[\mathrm{Var}_Y(Y \mid X)] + \mathrm{Var}_X(E_Y[Y \mid X]), \quad \text{where } X \text{ and } Y \text{ are random variables.} \qquad (2.4.2)$$

Letting Y be the pure premium PP and X be the number of claims N in the above formula gives:

$$\begin{aligned}\mathrm{Var}(PP) &= E_N[\mathrm{Var}_{PP}(PP \mid N)] + \mathrm{Var}_N(E_{PP}[PP \mid N])\\ &= E_N[N\sigma_S^2] + \mathrm{Var}_N(\mu_S N) = E_N[N]\sigma_S^2 + \mu_S^2\,\mathrm{Var}_N(N)\\ &= \mu_f\sigma_S^2 + \mu_S^2\sigma_f^2\end{aligned}$$

Where we have used the assumption that the frequency and severity are independent and the facts:

- For a fixed number of claims N, the variance of the pure premium is the variance of the sum of N independent identically distributed variables each with variance σ_S^2. (Since frequency and severity are assumed independent, σ_S^2 is the same for each value of N.) Such variances add so that $\mathrm{Var}_{PP}(PP \mid N) = N\sigma_S^2$.
- For a fixed number of claims N with frequency and severity independent, the expected value of the pure premium is N times the mean severity: $E_{PP}[PP \mid N] = N\mu_S$.

[9]The total variance = expected value of the process variance + the variation of the hypothetical means.

- Since with respect to N the variance of the severity acts as a constant:
$$E_N[N\sigma_S^2] = \sigma_S^2 E_N[N] = \mu_f \sigma_S^2$$

- Since with respect to N the mean of the severity acts as a constant:
$$\mathrm{Var}_N(\mu_S N) = \mu_S^2 \mathrm{Var}_N(N) = \mu_S^2 \sigma_f^2$$

Poisson Frequency

In the case of a Poisson Frequency with independent frequency and severity the formula for the process variance of the pure premium simplifies. Since $\mu_f = \sigma_f^2$:

$$\begin{aligned}\sigma_{PP}^2 &= \mu_f \sigma_S^2 + \mu_S^2 \sigma_f^2 \\ &= \mu_f(\sigma_S^2 + \mu_S^2) = \mu_f(\text{2nd moment of the severity})\end{aligned} \tag{2.4.3}$$

Example 2.4.4: Assume the following:

- For a given large risk, the number of claims for a single exposure period is Poisson with mean 3,645.
- The severity distribution is LogNormal with parameters $\mu = 5$ and $\sigma = 1.5$.
- Frequency and severity are independent.

Determine the variance of the pure premium for this risk.

[*Solution:* The second moment of the severity = $\exp(2\mu + 2\sigma^2)$ = $\exp(14.5)$ = 1,982,759.264. (See Appendix.) Thus $\sigma_{PP}^2 = \mu_f$(2nd moment of the severity) = (3,645)(1,982,759) = 7.22716×10^9.]

Normal Approximation:

For large numbers of expected claims, the observed pure premiums are approximately normally distributed.[10] For ex-

[10]The more skewed the severity distribution, the higher the frequency has to be for the Normal Approximation to produce worthwhile results.

ample, continuing the example above, mean severity = $\exp(\mu + .5\sigma^2) = \exp(6.125) = 457.14$. Thus the mean pure premium is $(3,645)(457.14) = 1,666,292$. One could ask what the chance is of the observed pure premiums being between 1.4997 million and 1.8329 million. Since the variance is 7.22716×10^9, the standard deviation of the pure premium is 85,013. Thus this probability of the observed pure premiums being within $\pm 10\%$ of 1.6663 million is

$$\begin{aligned}&\simeq \Phi((1.8329 \text{ million} - 1.6663 \text{ million})/85,013)\\ &\qquad - \Phi((1.4997 \text{ million} - 1.6663 \text{ million})/85,013)\\ &= \Phi(1.96) - \Phi(-1.96) = .975 - (1 - .975) = 95\%.\end{aligned}$$

Thus in this case with an expected number of claims equal to 3,645, there is about a 95% chance that the observed pure premium will be within $\pm 10\%$ of the expected value. One could turn this around and ask how many claims one would need in order to have a 95% chance that the observed pure premium will be within $\pm 10\%$ of the expected value. The answer of 3,645 claims could be taken as a Standard for Full Credibility for the Pure Premium.[11]

2.4. Exercises

2.4.1. Assume the following for a given risk:

- Mean frequency = 13; Variance of the frequency = 37
- Mean severity = 300; Variance of the severity = 200,000
- Frequency and severity are independent

What is the variance of the pure premium for this risk?

2.4.2. A six-sided die is used to determine whether or not there is a claim. Each side of the die is marked with either a 0 or a 1, where 0 represents no claim and 1 represents a

[11] As discussed in a subsequent section.

claim. Two sides are marked with a zero and four sides with a 1. In addition, there is a spinner representing claim severity. The spinner has three areas marked 2, 5 and 14. The probabilities for each claim size are:

Claim Size	Probability
2	20%
5	50%
14	30%

The die is rolled and if a claim occurs, the spinner is spun. What is the variance for a single trial of this risk process?

2.4.3. You are given the following:

- For a given risk, the number of claims for a single exposure period will be 1, with probability $4/5$; or 2, with probability $1/5$.
- If only one claim is incurred, the size of the claim will be 50, with probability $3/4$; or 200, with probability $1/4$.
- If two claims are incurred, the size of each claim, independent of the other, will be 50, with probability 60%; or 150, with probability 40%.

Determine the variance of the pure premium for this risk.

2.4.4. You are given the following:

- Number of claims for a single insured follows a Poisson distribution with mean .25
- The amount of a single claim has a uniform distribution on $[0, 5{,}000]$
- Number of claims and claim severity are independent.

Determine the pure premium's process variance for a single insured.

2.4.5. Assume the following:

- For the State of West Dakota, the number of claims for a single year is Poisson with mean 8,200
- The severity distribution is LogNormal with parameters $\mu = 4$ and $\sigma = 0.8$
- Frequency and severity are independent

Determine the expected aggregate losses. Determine the variance of the aggregate losses.

2.4.6. The frequency distribution follows the Poisson process with mean 0.5. The second moment about the origin for the severity distribution is 1,000. What is the process variance of the aggregate claim amount?

2.4.7. The probability function of claims per year for an individual risk is Poisson with a mean of 0.10. There are four types of claims. The number of claims has a Poisson distribution for each type of claim. The table below describes the characteristics of the four types of claims.

Type of Claim	Mean Frequency	Severity Mean	Severity Variance
W	.02	200	2,500
X	.03	1,000	1,000,000
Y	.04	100	0
Z	.01	1,500	2,000,000

Calculate the variance of the pure premium.

2.5. Full Credibility for Aggregate Losses, Pure Premiums, and Loss Ratios

Since they depend on both the number of claims and the size of claims, aggregate losses, pure premiums, and loss ratios have more reasons to vary than either frequency or severity. Because they are more difficult to estimate than frequencies, all other things being equal, the Standard for Full Credibility is larger than that for frequencies.

In Section 2.4 formulas for the variance of the pure premium were calculated:

General case: $$\sigma_{PP}^2 = \mu_f \sigma_S^2 + \mu_S^2 \sigma_f^2 \qquad (2.5.1)$$

Poisson frequency: $$\sigma_{PP}^2 = \mu_f(\sigma_S^2 + \mu_S^2) = \mu_f(\text{2nd moment of the severity}) \qquad (2.5.2)$$

The subscripts indicate the means and variances of the frequency (f) and severity (S). Assuming the Normal Approximation, full credibility standards can be calculated following the same steps as in Sections 2.2 and 2.3.

The probability that the observed pure premium PP is within $\pm k\%$ of the mean μ_{PP} is:

$$\begin{aligned} P &= \text{Prob}[\mu_{PP} - k\mu_{PP} \le PP \le \mu_{PP} + k\mu_{PP}] \\ &= \text{Prob}[-k(\mu_{PP}/\sigma_{PP}) \le u \le k(\mu_{PP}/\sigma_{PP})], \end{aligned}$$

where $u = (PP - \mu_{PP})/\sigma_{PP}$ is a unit normal variable, assuming the Normal Approximation.

Define y such that $\Phi(y) = (1 + P)/2$. (See Section 2.2 for more details.) Then, in order to have probability P that the observed pure premium will differ from the true pure premium by less than $\pm k\mu_{PP}$:

$$y = k(\mu_{PP}/\sigma_{PP}) \qquad (2.5.3)$$

To proceed further with formula (2.5.1) we need to know something about the frequency distribution function.

Suppose that frequency is a Poisson process and that n_F is the expected number of claims required for Full Credibility of the Pure Premium. Given n_F is the expected number of claims, then $\mu_f = \sigma_f^2 = n_F$ and, assuming frequency and severity are independent:

$$\mu_{PP} = \mu_f \mu_S = n_F \mu_S$$

and,

$$\sigma_{PP}^2 = \mu_f(\sigma_S^2 + \mu_S^2) = n_F(\sigma_S^2 + \mu_S^2).$$

Substituting for μ_{PP} and σ_{PP} in formula (2.5.3) gives:

$$y = k(n_F \mu_S / (n_F(\sigma_S^2 + \mu_S^2))^{1/2}).$$

Solving for n_F:

$$n_F = (y/k)^2[1 + (\sigma_S^2/\mu_S^2)] = n_0(1 + CV_S^2) \qquad (2.5.4)$$

This is the **Standard for Full Credibility of the Pure Premium.** $n_0 = (y/k)^2$ is the Standard for Full Credibility of Frequency that was derived in Section 2.2. $CV_S = (\sigma_S/\mu_S)$ is the coefficient of variation of the severity. Formula (2.5.4) can also be written as $n_F = n_0(\mu_S^2 + \sigma_S^2)/\mu_S^2$ where $(\mu_S^2 + \sigma_S^2)$ is the second moment of the severity distribution.

Example 2.5.1: The number of claims has a Poisson distribution. The mean of the severity distribution is 2,000 and the standard deviation is 4,000. For $P = 90\%$ and $k = 5\%$, what is the Standard for Full Credibility of the Pure Premium?

[*Solution:* From section 2.2, $n_0 = 1{,}082$ claims. The coefficient of variation is $CV = 4{,}000/2{,}000 = 2$. So, $n_F = 1{,}082$ $(1 + 2^2) = 5{,}410$ claims.]

It is interesting to note that the Standard for Full Credibility of the Pure Premium is the sum of the standards for frequency and severity:

$$\begin{aligned} n_F &= n_0(1 + CV_S^2) = n_0 + n_0 CV_S^2 \\ &= \text{Standard for Full Credibility of Frequency} \\ &\quad + \text{Standard for Full Credibility of Severity} \end{aligned}$$

Note that if one limits the size of claims, then the coefficient of variation is smaller. Therefore, the criterion for full credibility for basic limits losses is less than that for total losses. It is a common practice in ratemaking to cap losses in order to increase the credibility assigned to the data.

The pure premiums are often approximately Normal; generally the greater the expected number of claims or the shorter tailed the frequency and severity distributions, the better the Normal Approximation. It is assumed that one has enough claims that the aggregate losses approximate a Normal Distribution. While it is possible to derive formulas that don't depend on the Normal Approximation, they're not covered here.[12]

Variations from the Poisson Assumption

As with the Standard for Full Credibility of Frequency, one can derive a more general formula when the Poisson assumption does not apply. The Standard for Full Credibility is:[13]

$$n_F = \{y^2/k^2\}(\sigma_f^2/\mu_f + \sigma_s^2/\mu_s^2), \tag{2.5.5}$$

which reduces to the Poisson case when $\sigma_f^2/\mu_f = 1$. If the severity is constant then σ_s^2 is zero and (2.5.5) reduces to (2.2.6).

2.5. Exercises

[*Assume that frequency and severity are independent in the following problems*.]

[12]One can, for example, use the Normal Power Approximation, which takes into account more than the first two moments. See for example, "Limited Fluctuation Credibility with the Normal Power Approximation" by Gary Venter. This usually has little practical effect.
[13]A derivation can be found in Mayerson, et al, "The Credibility of the Pure Premium."

2.5.1. You are given the following information:

- The number of claims is Poisson.
- The severity distribution is LogNormal with parameters $\mu = 4$ and $\sigma = 0.8$.
- Full credibility is defined as having a 90% probability of being within plus or minus 2.5% of the true pure premium.

What is the minimum number of expected claims that will be given full credibility?

2.5.2. Given the following information, what is the minimum number of *policies* that will be given full credibility?

- Mean claim frequency = .04 claims per policy. (Assume Poisson.)
- Mean claim severity = \$1,000.
- Variance of the claim severity = \$2 million.
- Full credibility is defined as having a 99% probability of being within plus or minus 10% of the true pure premium.

2.5.3. The full credibility standard for a company is set so that the total number of claims is to be within 2.5% of the true value with probability P. This full credibility standard is calculated to be 5,000 claims. The standard is altered so that the total *cost* of claims is to be within 9% of the true value with probability P. The claim frequency has a Poisson distribution and the claim severity has the following distribution:

$$f(x) = .0008(50 - x), \qquad 0 \leq x \leq 50.$$

What is the expected number of claims necessary to obtain full credibility under the new standard?

2.5.4. You are given the following information:

- A standard for full credibility of 2,000 claims has been selected so that the actual pure premium would be within 10% of the expected pure premium 99% of the time.
- The number of claims follows a Poisson distribution.

Using the classical credibility concepts determine the coefficient of variation of the severity distribution underlying the full credibility standard.

2.5.5. You are given the following:

- The number of claims is Poisson distributed.
- Claim severity has the following distribution:

Claim Size	Probability
10	.50
20	.30
50	.20

Determine the number of claims needed so that the total cost of claims is within 20% of the expected cost with 95% probability.

2.5.6. You are given the following:

- The number of claims has a *negative binomial* distribution with a variance that is twice as large as the mean.
- Claim severity has the following distribution:

Claim Size	Probability
10	.50
20	.30
50	.20

Determine the number of claims needed so that the total cost of claims is within 20% of the expected cost with 95% probability. Compare your answer to that of exercise 2.5.5.

2.5.7. A full credibility standard is determined so that the total *number* of claims is within 2.5% of the expected number with probability 98%. If the same expected number of claims for full credibility is applied to the total *cost* of claims, the actual total cost would be within $100k\%$ of the expected cost with 90% probability. The coefficient of variation of the severity is 3.5. The frequency is Poisson. Using the normal approximation of the aggregate loss distribution, determine k.

2.5.8. The ABC Insurance Company has decided to use Classical Credibility methods to establish its credibility requirements for an individual state rate filing. The full credibility standard is to be set so that the observed total cost of claims underlying the rate filing should be within 5% of the true value with probability 0.95. The claim frequency follows a Poisson distribution and the claim severity is distributed according to the following distribution:

$$f(x) = 1/100{,}000 \qquad \text{for} \quad 0 \le x \le 100{,}000$$

What is the expected number of claims, N_F necessary to obtain full credibility?

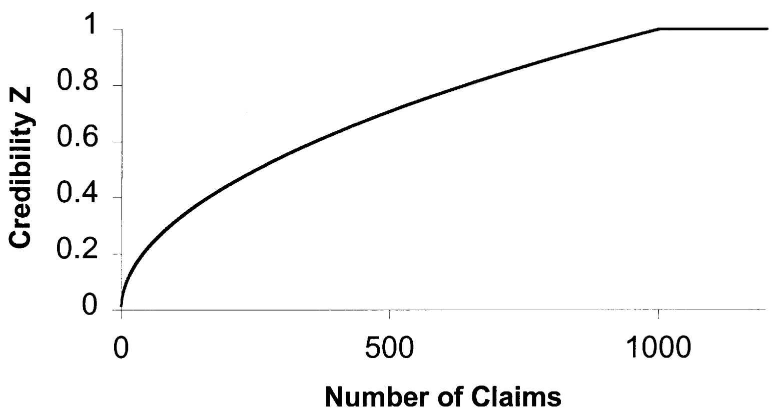
2.5.9. A full credibility standard of 1,200 expected claims has been established for aggregate claim costs. Determine the number of expected claims that would be required for full credibility if the coefficient of variation of the claim size distribution were changed from 2 to 4 and the range parameter, k, were doubled.

2.6. Partial Credibility

When one has at least the number of claims needed for Full Credibility, then one assigns 100% credibility to the observa-

GRAPH 1

CLASSICAL CREDIBILITY



tions. However, when there is less data than is needed for full credibility, less that 100% credibility is assigned.

Let n be the (expected) number of claims for the volume of data, and n_F be the standard for Full Credibility. Then the partial credibility assigned is $Z = \sqrt{n/n_F}$. If $n \geq n_F$, then $Z = 1.00$. Use the square root rule for partial credibility for either frequency, severity or pure premiums.

For example if 1,000 claims are needed for full credibility, then Graph 1 displays the credibilities that would be assigned.

Example 2.6.1: The Standard for Full Credibility is 683 claims and one has observed 300 claims.[14] How much credibility is assigned to this data?

[14]Ideally, n in the formula $Z = \sqrt{n/n_F}$ should be the expected number of claims. However, this is often not known and the observed number of claims is used as an approximation. If the number of exposures is known along with an expected claims frequency, then the expected number of claims can be calculated by (number of exposures) $\times$ (expected claims frequency).

[*Solution:* $\sqrt{300/683} = 66.3\%$.]

Limiting Fluctuations

The square root rule for partial credibility is designed so that the standard deviation of the contribution of the data to the new estimate retains the value corresponding to the standard for full credibility. We will demonstrate why the square root rule accomplishes that goal. One does not need to follow the derivation in order to apply the simple square root rule.

Let X_{partial} be a value calculated from partially credible data; for example, X_{partial} might be the claim frequency calculated from the data. Assume X_{full} is calculated from data that just meets the full credibility standard. For the full credibility data, *Estimate* = X_{full}, while the partially credible data enters the estimate with a weight Z in front of it: *Estimate* $= ZX_{\text{partial}} + (1 - Z)$[*Other Information*]. The credibility Z is calculated so that the expected variation in ZX_{partial} is limited to the variation allowed in a full credibility estimate X_{full}. The variance of ZX_{partial} can be reduced by choosing a Z less than one.

Suppose you are trying to estimate frequency (number of claims per exposure), pure premium, or loss ratio, with estimates X_{partial} and X_{full} based on different size samples of a population. Then, they will have the same expected value μ. But, since it is based on a smaller sample size, X_{partial} will have a larger standard deviation σ_{partial} than the standard deviation σ_{full} of the full credibility estimate X_{full}. The goal is to limit the fluctuation in the term ZX_{partial} to that allowed for X_{full}. This can be written as:[15]

$$\begin{aligned}\text{Prob}[\mu - k\mu \leq X_{\text{full}} \leq \mu + k\mu]\\ = \text{Prob}[Z\mu - k\mu \leq ZX_{\text{partial}} \leq Z\mu + k\mu]\end{aligned}$$

[15]Note that in both cases fluctuations are limited to $\pm k\mu$ of the mean.

Subtracting through by the means and dividing by the standard deviations gives:

$$\begin{aligned}\text{Prob}[-k\mu/\sigma_{\text{full}} &\leq (X_{\text{full}} - \mu)/\sigma_{\text{full}} \leq k\mu/\sigma_{\text{full}}] \\ &= \text{Prob}[-k\mu/Z\sigma_{\text{partial}} \leq (ZX_{\text{partial}} - Z\mu)/ \\ &\qquad Z\sigma_{\text{partial}} \leq k\mu/Z\sigma_{\text{partial}}]^{16}\end{aligned}$$

Assuming the Normal Approximation, $(X_{\text{full}} - \mu)/\sigma_{\text{full}}$ and $(ZX_{\text{partial}} - Z\mu)/Z\sigma_{\text{partial}}$ are unit normal variables. Then, the two sides of the equation are equal if:

$$k\mu/\sigma_{\text{full}} = k\mu/Z\sigma_{\text{partial}}$$

Solving for Z yields:

$$Z = \sigma_{\text{full}}/\sigma_{\text{partial}}. \tag{2.6.1}$$

Thus the partial credibility Z will be inversely proportional to the standard deviation of the partially credible data.

Assume we are trying to estimate the average number of accidents μ in a year per driver for a homogeneous population. For a sample of M drivers, $\mu_M = \sum_{i=1}^{M} m_i/M$ is an estimate of the frequency μ where m_i is the number of accidents for the i^{th} driver. Assuming that the numbers of claims per driver are independent of each other, then the variance of μ_M is $\sigma_M^2 = \text{Var}[\sum_{i=1}^{M} m_i/M] = (1/M^2)\sum_{i=1}^{M} \text{Var}(m_i)$. If each insured has a Poisson frequency with the same mean $\mu = \text{Var}(m_i)$, then $\sigma_M^2 = (1/M^2)\sum_{i=1}^{M} \mu = M\mu/M^2 = \mu/M$.

If a sample of size M is expected to produce n claims, then since $M\mu = n$, it follows that $M = n/\mu$. So, the variance is $\sigma_M^2 = \mu/M = \mu/(n/\mu) = \mu^2/n$, and the standard deviation is:

$$\sigma_M = \mu/\sqrt{n}. \tag{2.6.2}$$

Example 2.6.2: A sample with 1,000 expected claims is used to estimate frequency μ. Assuming frequency is Poisson, what

[16] Note that the mean of ZX_{partial} is $Z\mu$ and the standard deviation is $Z\sigma_{\text{partial}}$.

are the variance and standard deviation of the estimated frequency?

[*Solution:* The variance is $\mu^2/1000$ and the standard deviation is $\mu/\sqrt{1{,}000} = .032\mu$.]

A fully credible sample with an expected number of claims n_0, will have a standard deviation $\sigma_{\text{full}} = \mu/\sqrt{n_0}$. A partially credible sample with expected number of claims n will have $\sigma_{\text{partial}} = \mu/\sqrt{n}$. Using formula (2.6.1), the credibility for the smaller sample is: $Z = (\mu/\sqrt{n_0})/(\mu/\sqrt{n}) = \sqrt{n/n_0}$. So,

$$Z = \sqrt{n/n_0} \tag{2.6.3}$$

Equation 2.6.3 is the important square root rule for partial credibility. Note that the Normal Approximation and Poisson claims distribution were assumed along the way. A similar derivation of the square root formula also applies to credibility for severity and the pure premium.[17]

2.6. Exercises

2.6.1. The Standard for Full Credibility is 2,000 claims. How much credibility is assigned to 300 claims?

2.6.2. Using the square root rule for partial credibility, a certain volume of data is assigned credibility of .36. How much credibility would be assigned to ten times that volume of data?

2.6.3. Assume a Standard for Full Credibility for *severity* of 2,500 claims. For the class of Salespersons one has observed 803 claims totaling \$9,771,000. Assume the av-

[17]The square root formula for partial credibility also applies in the calculation of aggregate losses and total number of claims although equation (2.6.1) needs to be revised. For estimates of aggregate losses and total number of claims, a larger sample will have a larger standard deviation. Letting $L = X_1 + X_2 + \cdots + X_N$ represent aggregate losses, then the standard deviation of L increases as the number of expected claims increases, but the ratio of the standard deviation of L to the expected value of L decreases. Equation (2.6.1) will work if the standard deviations are replaced by coefficients of variation.

erage cost per claim for all similar classes is \$10,300. Calculate a credibility-weighted estimate of the average cost per claim for the Salespersons' class.

2.6.4. The Standard for Full Credibility is 3,300 claims. The expected claim frequency is 6% per house-year. How much credibility is assigned to 2,000 house-years of data?

2.6.5. You are given the following information:

- Frequency is Poisson.
- Severity follows a Gamma Distribution with $\alpha = 1.5$.
- Frequency and severity are independent.
- Full credibility is defined as having a 97% probability of being within plus or minus 4% of the true pure premium.

What credibility is assigned to 150 claims?

2.6.6. The 1984 pure premium underlying the rate equals \$1,000. The loss experience is such that the observed pure premium for that year equals \$1,200 and the number of claims equals 600. If 5,400 claims are needed for full credibility and the square root rule for partial credibility is used, estimate the pure premium underlying the rate in 1985. (Assume no change in the pure premium due to inflation.)

2.6.7. Assume the random variable N, representing the number of claims for a given insurance portfolio during a one-year period, has a Poisson distribution with a mean of n. Also assume $X_1, X_2 \ldots, X_N$ are N independent, identically distributed random variables with X_i representing the size of the i^{th} claim. Let $C = X_1 + X_2 + \cdots X_n$ represent the total cost of claims during a year. We want to use the observed value of C as an estimate of future costs. We are willing to assign full credibility to C provided it is within

10.0% of its expected value with probability 0.96. If the claim size distribution has a coefficient of variation of 0.60, what credibility should we assign to the experience if 213 claims occur?

2.6.8. The Slippery Rock Insurance Company is reviewing their rates. The expected number of claims necessary for full credibility is to be determined so that the observed total cost of claims should be within 5% of the true value 90% of the time. Based on independent studies, they have estimated that individual claims are independently and identically distributed as follows:

$$f(x) = 1/200{,}000, \qquad 0 \leq x \leq 200{,}000.$$

Assume that the number of claims follows a Poisson distribution. What is the credibility Z to be assigned to the most recent experience given that it contains 1,082 claims?

2.6.9. You are given the following information for a group of insureds:

- Prior estimate of expected total losses $20,000,000
- Observed total losses $25,000,000
- Observed number of claims 10,000
- Required number of claims for full credibility 17,500

Calculate a credibility weighted estimate of the group's expected total losses.

2.6.10. 2,000 expected claims are needed for full credibility. Determine the number of expected claims needed for 60% credibility.

2.6.11. The full credibility standard has been selected so that the actual *number of claims* will be within 5% of the expected *number of claims* 90% of the time. Determine the credibility to be given to the experience if 500 claims are expected.

3. LEAST SQUARES CREDIBILITY

The second form of credibility covered in this chapter is called Least Squares or Bühlmann Credibility. It is also referred as greatest accuracy credibility. As will be discussed, the credibility is given by the formula: $Z = N/(N + K)$. As the number of observations N increases, the credibility Z approaches 1.

In order to apply Bühlmann Credibility to various real-world situations, one is typically required to calculate or estimate the so-called Bühlmann Credibility Parameter K. This involves being able to apply analysis of variance: the calculation of the expected value of the process variance and the variance of the hypothetical means.

Therefore, in this section we will first cover the calculation of the expected value of the process variance and the variance of the hypothetical means. This will be followed by applications of Bühlmann Credibility to various simplified situations. Finally, we will illustrate the ideas covered via the excellent Philbrick Target Shooting Example.

3.1. Analysis of Variance

Let's start with an example involving multi-sided dice:

There are a total of 100 multi-sided dice of which 60 are 4-sided, 30 are 6-sided and 10 are 8-sided. The multi-sided dice with 4 sides have 1, 2, 3 and 4 on them. The multi-sided dice with the usual 6 sides have numbers 1 through 6 on them. The multi-sided dice with 8 sides have numbers 1 through 8 on them. For a given die each side has an equal chance of being rolled; i.e., the die is fair.

Your friend picked at random a multi-sided die. He then rolled the die and told you the result. You are to estimate the result when he rolls that same die again.

The next section will demonstrate how to apply Bühlmann Credibility to this problem. In order to apply Bühlmann Credibility one will first have to calculate the items that would be used in "analysis of variance." One needs to compute the Expected Value of the Process Variance and the Variance of the Hypothetical Means, which together sum to the total variance.

Expected Value of the Process Variance:
For each type of die we can compute the mean and the (process) variance. For example, for a 6-sided die one need only list all the possibilities:

A	B	C	D
Roll of Die	A Priori Probability	Column A× Column B	Square of Column A ×Column B
1	0.16667	0.16667	0.16667
2	0.16667	0.33333	0.66667
3	0.16667	0.50000	1.50000
4	0.16667	0.66667	2.66667
5	0.16667	0.83333	4.16667
6	0.16667	1.00000	6.00000
Sum	1	3.5	15.16667

Thus the mean is 3.5 and the variance is $15.16667 - 3.5^2 = 2.91667 = 35/12$. Thus the conditional variance if a 6-sided die is picked is: $\mathrm{Var}[X \mid 6\text{-sided}] = 35/12$.

Example 3.1.1: What is the mean and variance of a 4-sided die?

[*Solution:* The mean is 2.5 and the variance is 15/12.]

Example 3.1.2: What is the mean and variance of an 8-sided die?

[*Solution:* The mean is 4.5 and the variance is 63/12.]

One computes the **Expected Value of the Process Variance (EPV)** by weighting together the process variances for each type of risk using as weights the chance of having each type of risk.[18] In this case the Expected Value of the Process Variance is: $(60\%)(15/12) + (30\%)(35/12) + (10\%)(63/12) = 25.8/12 = 2.15$. In symbols this sum is: $P(\text{4-sided})\,\text{Var}[X \mid \text{4-sided}] + P(\text{6-sided})\,\text{Var}[X \mid \text{6-sided}] + P(\text{8-sided})\,\text{Var}[X \mid \text{8-sided}]$. Note that this is the Expected Value of the Process Variance for *one* observation of the risk process; i.e., one roll of a die.

Variance of the Hypothetical Means

The hypothetical means are 2.5, 3.5, and 4.5 for the 4-sided, 6-sided, and 8-sided die, respectively. One can compute the **Variance of the Hypothetical Means (VHM)** by the usual technique; compute the first and second moments of the hypothetical means.

Type of Die	A Priori[19] Chance of this Type of Die	Mean for this Type of Die	Square of Mean of this Type of Die
4-sided	0.6	2.5	6.25
6-sided	0.3	3.5	12.25
8-sided	0.1	4.5	20.25
Average		3	9.45

The Variance of the Hypothetical Means is the second moment minus the square of the (overall) mean $= 9.45 - 3^2 = .45$. Note

[18]In situations where the types of risks are parametrized by a continuous distribution, as for example in the Gamma-Poisson frequency process, one will take an integral rather than a sum.

[19]According to the dictionary, a priori means "relating to or derived by reasoning from self-evident propositions." This usage applies here since we can derive the probabilities from the statement of the problem. After we observe rolls of the die, we may calculate new probabilities that recognize both the a priori values and the observations. This is covered in detail in section 4.

that this is the variance for a single observation, i.e., one roll of a die.

Total Variance

One can compute the total variance of the observed results if one were to do this experiment repeatedly. One needs merely compute the chance of each possible outcome.

In this case there is a $60\% \times (1/4) = 15\%$ chance that a 4-sided die will be picked and then a 1 will be rolled. Similarly, there is a $30\% \times (1/6) = 5\%$ chance that a 6-sided die will be selected and then a 1 will be rolled. There is a $10\% \times (1/8) = 1.25\%$ chance that an 8-sided die will be selected and then a 1 will be rolled. The total chance of a 1 is therefore:

$$15\% + 5\% + 1.25\% = 21.25\%.$$

A	B	C	D	E	F	G
Roll of Die	Probability due to 4-sided die	Probability due to 6-sided die	Probability due to 8-sided die	A Priori Probability $= B + C + D$	Column A $\times$Column E	Square of Column A $\times$Column E
1	0.15	0.05	0.0125	0.2125	0.2125	0.2125
2	0.15	0.05	0.0125	0.2125	0.4250	0.8500
3	0.15	0.05	0.0125	0.2125	0.6375	1.9125
4	0.15	0.05	0.0125	0.2125	0.8500	3.4000
5		0.05	0.0125	0.0625	0.3125	1.5625
6		0.05	0.0125	0.0625	0.3750	2.2500
7			0.0125	0.0125	0.0875	0.6125
8			0.0125	0.0125	0.1000	0.8000
Sum	0.6	0.3	0.1	1	3	11.6

The mean is 3 (the same as computed above) and the second moment is 11.6. Therefore, the total variance is $11.6 - 3^2 = 2.6$. Note that Expected Value of the Process Variance + Variance of the Hypothetical Means = 2.15 + .45 = 2.6 = Total Variance. Thus the total variance has been split into two pieces. This is true in general.

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

(3.1.1)

While the two pieces of the total variance seem similar, the order of operations in their computation is different. In the case of the Expected Value of the Process Variance, EPV, first one separately computes the process variance for each of the types of risks and then one takes the expected value over all types of risks. Symbolically, the **EPV** = $E_\theta[\mathrm{Var}[X \mid \theta]]$.

In the case of the Variance of the Hypothetical Means, VHM, first one computes the expected value for each type of risk and then one takes their variance over all types of risks. Symbolically, the **VHM** = $\mathrm{Var}_\theta[E[X \mid \theta]]$.

Multiple Die Rolls

So far we have computed variances in the case of a single roll of a die. One can also compute variances when one is rolling more than one die.[20] There are a number of somewhat different situations which lead to different variances, which lead in turn to different credibilities.

Example 3.1.3: Each actuary attending a CAS Meeting rolls two multi-sided dice. One die is 4-sided and the other is 6-sided. Each actuary rolls his two dice and reports the sum. What is the expected variance of the results reported by the actuaries?

[*Solution:* The variance is the sum of that for a 4-sided and 6-sided die. Variance = $(15/12) + (35/12) = 50/12 = 4.167$.]

One has to distinguish the situation in example 3.1.3 where the types of dice rolled are known, from one where each actuary is selecting dice at random. The latter introduces an additional source of random variation, as shown in the following exercise.

[20]These dice examples can help one to think about insurance situations where one has more than one observation or insureds of different sizes.

Example 3.1.4: Each actuary attending a CAS Meeting independently selects two multi-sided dice. For each actuary his two multi-sided dice are selected *independently* of each other, with each die having a 60% chance of being 4-sided, a 30% chance of being 6-sided, and a 10% chance of being 8-sided. Each actuary rolls his two dice and reports the sum. What is the expected variance of the results reported by the actuaries?

[*Solution:* The total variance is the sum of the EPV and VHM. For each actuary let his two dice be A and B. Let the parameter (number of sides) for A be θ and that for B be ψ. Note that A only depends on θ, while B only depends on ψ, since the two dice were selected independently. Then EPV $= E_{\theta,\psi}[\mathrm{Var}[A+B \mid \theta,\psi]] = E_{\theta,\psi}[\mathrm{Var}[A \mid \theta,\psi]] + E_{\theta,\psi}[\mathrm{Var}[B \mid \theta,\psi]] = E_{\theta}[\mathrm{Var}[A \mid \theta]] + E_{\psi}[\mathrm{Var}[B \mid \psi]] = 2.15 + 2.15 = (2)(2.15) = 4.30$. The VHM $= \mathrm{Var}_{\theta,\psi}[E[A+B \mid \theta,\psi]] = \mathrm{Var}_{\theta,\psi}[E[A \mid \theta,\psi] + E[B \mid \theta,\psi]] = \mathrm{Var}_{\theta}[E[A \mid \theta]] + \mathrm{Var}_{\psi}[E[B \mid \psi]] = (2)(.45) = .90$. Where we have used the fact that $E[A \mid \theta]$ and $E[B \mid \psi]$ are independent and thus their variances add. Total variance = EPV + VHM = $4.3 + .9 = 5.2$.]

Example 3.1.4 is subtly different from a situation where the two dice selected by a given actuary are always of the same type, as in example 3.1.5.

Example 3.1.5: Each actuary attending a CAS Meeting selects two multi-sided dice both of the same type. For each actuary, his multi-sided dice have a 60% chance of being 4-sided, a 30% chance of being 6-sided, and a 10% chance of being 8-sided. Each actuary rolls his dice and reports the sum. What is the expected variance of the results reported by the actuaries?

[*Solution:* The total variance is the sum of the EPV and VHM. For each actuary let his two die rolls be A and B. Let the parameter (number of sides) for his dice be θ, the same for both dice. Then EPV $= E_{\theta}[\mathrm{Var}[A+B \mid \theta]] = E_{\theta}[\mathrm{Var}[A \mid \theta]] + E_{\theta}[\mathrm{Var}[B \mid \theta]] = E_{\theta}[\mathrm{Var}[A \mid \theta]] + E_{\theta}[\mathrm{Var}[B \mid \theta]] = 2.15 + 2.15 = (2)(2.15) = 4.30$. The VHM $= \mathrm{Var}_{\theta}[E[A+B \mid \theta]] = \mathrm{Var}_{\theta}[2E[A \mid \theta]] = (2^2)\,\mathrm{Var}_{\theta}[E[A \mid \theta]] = (4)(.45) = 1.80$. Where we have used the fact that $E[A \mid \theta]$ and $E[B \mid \theta]$ are the same. So, Total Variance = EPV + VHM = 4.3 + 1.8 = 6.1. Alternately, Total Variance = (N)(EPV for one observation) + (N^2)(VHM for one observation) = (2)(2.15) + (2^2)(.45) = 6.1.]

Note that example 3.1.5 is the same mathematically as if each actuary chose a single die and reported the sum of rolling his die twice. Contrast this with previous example 3.1.4 in which each actuary chose two dice, with the type of each die independent of the other.

In example 3.1.5: Total Variance = (2)(EPV single die roll) +(2^2)(VHM single die roll).

The VHM has increased in proportion to N^2, the square of the number of observations, while the EPV goes up only as N:

$$\text{Total Variance} = N(\text{EPV for one observation}) + (N^2)(\text{VHM for one observation}) \tag{3.1.2}$$

This is the assumption behind the Bühlmann Credibility formula: $Z = N/(N + K)$. The Bühlmann Credibility parameter K is the ratio of the EPV to the VHM for a *single* die. The formula automatically adjusts the credibility for the number of observations N.

Total Variance = EPV + VHM

One can demonstrate that in general:

$$\text{Var}[X] = E_\theta[\text{Var}[X \mid \theta]] + \text{Var}_\theta[E[X \mid \theta]]$$

First one can rewrite the EPV: $E_\theta[\text{Var}[X \mid \theta]] = E_\theta[E[X^2 \mid \theta] - E[X \mid \theta]^2] = E_\theta[E[X^2 \mid \theta]] - E_\theta[E[X \mid \theta]^2] = E[X^2] - E_\theta[E[X \mid \theta]^2]$.

Second, one can rewrite the VHM: $\text{Var}_\theta[E[X \mid \theta]] = E_\theta[E[X \mid \theta]^2] - E_\theta[E[X \mid \theta]]^2 = E_\theta[E[X \mid \theta]^2] - E[X]^2 = E_\theta[E[X \mid \theta]^2] - E[X]^2$.

Putting together the first two steps: EPV + VHM = $E_\theta[\text{Var}[X \mid \theta]] + \text{Var}_\theta[E[X \mid \theta]] = E[X^2] - E_\theta[E[X \mid \theta]^2] + E_\theta[E[X \mid \theta]^2] - E[X]^2 = E[X^2] - E[X]^2 = \text{Var}[X]$ = Total Variance of X.

In the case of the single die: $2.15 + .45 = (11.6 - 9.45) + (9.45 - 9) = 11.6 - 9 = 2.6$. In order to split the total variance of 2.6 into two pieces we've added and subtracted the expected value of the squares of the hypothetical means: 9.45.

A Series of Examples

The following information will be used in a series of examples involving the frequency, severity, and pure premium:

Type	Portion of Risks in this Type	Bernoulli (Annual) Frequency Distribution[21]	Gamma Severity Distribution[22]
1	50%	$p = 40\%$	$\alpha = 4,\ \lambda = .01$
2	30%	$p = 70\%$	$\alpha = 3,\ \lambda = .01$
3	20%	$p = 80\%$	$\alpha = 2,\ \lambda = .01$

We assume that the types are homogeneous; i.e., every insured of a given type has the same frequency and severity process. Assume that for an individual insured, frequency and severity are independent.[23]

We will show how to compute the Expected Value of the Process Variance and the Variance of the Hypothetical Means in each case. In general, the simplest case involves the frequency, followed by the severity, with the pure premium being the most complex case.

Expected Value of the Process Variance, Frequency Example

For type 1, the process variance of the Bernoulli frequency is $pq = (.4)(1 - .4) = .24$. Similarly, for type 2 the process variance

[21] With a Bernoulli frequency distribution, the probability of exactly one claim is p and the probability of no claims is $(1 - p)$. The mean of the distribution is p and the variance is pq where $q = (1 - p)$.

[22] For the Gamma distribution, the mean is α/λ and the variance is α/λ^2. See the Appendix on claim frequency and severity distributions.

[23] Across types, the frequency and severity are not independent. In this example, types with higher average frequency have lower average severity.

for the frequency is $(.7)(1-.7) = .21$. For type 3 the process variance for the frequency is $(.8)(1-.8) = .16$.

The expected value of the process variance is the weighted average of the process variances for the individual types, using the a priori probabilities as the weights. The EPV of the frequency = $(50\%)(.24) + (30\%)(.21) + (20\%)(.16) = .215$.

Note that to compute the EPV one first computes variances and then computes the expected value. In contrast, in order to compute the VHM, one first computes expected values and then computes the variance.

Variance of the Hypothetical Mean Frequencies

For type 1, the mean of the Bernoulli frequency is $p = .4$. Similarly for type 2 the mean frequency is .7. For type 3 the mean frequency is .8.

The variance of the hypothetical mean frequencies is computed the same way any other variance is. First one computes the first moment: $(50\%)(.4) + (30\%)(.7) + (20\%)(.8) = .57$. Then one computes the second moment: $(50\%)(.4^2) + (30\%)(.7^2) + (20\%)(.8^2) = .355$. Then the VHM = $.355 - .57^2 = .0301$.

Expected Value of the Process Variance, Severity Example

The computation of the EPV for severity is similar to that for frequency with one important difference. One has to weight together the process variances of the severities for the individual types using the chance that a claim came from each type.[24] The chance that a claim came from an individual of a given type is proportional to the product of the a priori chance of an insured being of that type and the mean frequency for that type.

[24]Each claim is one observation of the severity process. The denominator for severity is number of claims. In contrast, the denominator for frequency (as well as pure premiums) is exposures.

For type 1, the process variance of the Gamma severity is $\alpha/\lambda^2 = 4/.01^2 = 40{,}000$. Similarly, for type 2 the process variance for the severity is $3/.01^2 = 30{,}000$. For type 3 the process variance for the severity is $2/.01^2 = 20{,}000$.

The mean frequencies are: .4, .7, and .8. The a priori chances of each type are: 50%, 30% and 20%. Thus the weights to use to compute the EPV of the severity are (.4)(50%) = .2, (.7)(30%) = .21, and (.8)(20%) = .16. The sum of the weights is .2 + .21 + .16 = .57. Thus the probability that a claim came from each class is: .351, .368, and .281. (For example, .2/.57 = .351.) The expected value of the process variance of the severity is the weighted average of the process variances for the individual types, using these weights.[25] The EPV of the severity[26] = {(.2)(40,000) + (.21)(30,000) + (.16)(20,000)}/(.2 + .21 + .16) = 30,702.

This computation can be organized in the form of a spreadsheet:

A	B	C	D	E	F	G	H
Class	A Priori Probability	Mean Frequency	Weights = Col. B ×Col. C	Probability that a Claim Came from this Class	Gamma Parameters α	λ	Process Variance
1	50%	0.4	0.20	0.351	4	0.01	40,000
2	30%	0.7	0.21	0.368	3	0.01	30,000
3	20%	0.8	0.16	0.281	2	0.01	20,000
Average			0.57	1.000			**30,702**

[25]Note that while in this case with discrete possibilities we take a sum, in the continuous case we would take an integral.

[26]Note that this result differs from what one would get by using the a priori probabilities as weights. The latter *incorrect* method would result in: (50%)(40,000) + (30%)(30,000) + (20%)(20,000) = 33,000 $\neq$ 30,702.

Variance of the Hypothetical Mean Severities

In computing the moments one again has to use for each individual type the chance that a claim came from that type.[27]

For type 1, the mean of the Gamma severity is $\alpha/\lambda = 4/.01 = 400$. Similarly for type 2 the mean severity is $3/.01 = 300$. For type 3 the mean severity is $2/.01 = 200$.

The mean frequencies are: .4, .7, and .8. The a priori chances of each type are: 50%, 30% and 20%. Thus the weights to use to compute the moments of the severity are (.4)(50%) = .2, (.7)(30%) = .21, and (.8)(20%) = .16.

The variance of the hypothetical mean severities is computed the same way any other variance is. First one computes the first moment: $\{(.2)(400) + (.21)(300) + (.16)(200)\}/(.2 + .21 + .16) = 307.02$. Then one computes the second moment: $\{(.2)(400^2) + (.21)(300^2) + (.16)(200^2)\}/(.2 + .21 + .16) = 100{,}526$. Then the VHM of the severity $= 100{,}526 - 307.02^2 = 6{,}265$. This computation can be organized in the form of a spreadsheet:[28]

A	B	C	D	E	F	G	H
Class	A Priori Probability	Mean Frequency	Weights = Col. B ×Col. C	Gamma Parameters α	λ	Mean Severity	Square of Mean Severity
1	50%	0.4	0.20	4	0.01	400	160,000
2	30%	0.7	0.21	3	0.01	300	90,000
3	20%	0.8	0.16	2	0.01	200	40,000
Average			0.57			307.02	100,526

[27] Each claim is one observation of the severity process. The denominator for severity is number of claims. In contrast, the denominator for frequency (as well as pure premiums) is exposures.

[28] After Column D, one could inset another column normalizing the weights by dividing them each by the sum of Column D. In the spreadsheet shown, one has to remember to divide by the sum of Column D when computing each of the moments.

Then the variance of the hypothetical mean severities = $100{,}526 - 307.02^2 = 6{,}265$.

Expected Value of the Process Variance, Pure Premium Example

The computation of the EPV for the pure premiums is similar to that for frequency. However, it is more complicated to compute each process variance of the pure premiums.

For type 1, the mean of the Bernoulli frequency is $p = .4$, and the variance of the Bernoulli frequency is $pq = (.4)(1 - .4) = .24$. For type 1, the mean of the Gamma severity is $\alpha/\lambda = 4/.01 = 400$, and the variance of the Gamma severity is $\alpha/\lambda^2 = 4/.01^2 = 40{,}000$. Thus since frequency and severity are assumed to be independent, the process variance of the pure premium = (Mean Frequency)(Variance of Severity) + (Mean Severity)^2 (Variance of Frequency) = $(.4)(40{,}000) + (400)^2(.24) = 54{,}400$.

Similarly for type 2 the process variance of the pure premium = $(.7)(30{,}000) + (300)^2(.21) = 39{,}900$. For type 3 the process variance of the pure premium = $(.8)(20{,}000) + (200)^2(.16) = 22{,}400$.

The expected value of the process variance is the weighted average of the process variances for the individual types, using the a priori probabilities as the weights. The EPV of the pure premium = (50%)(54,400) + (30%)(39,900) + (20%)(22,400) = 43,650. This computation can be organized in the form of a spreadsheet:

Class	A Priori Probability	Mean Frequency	Variance of Frequency	Mean Severity	Variance of Severity	Process Variance
1	50%	0.4	0.24	400	40,000	54,400
2	30%	0.7	0.21	300	30,000	39,900
3	20%	0.8	0.16	200	20,000	22,400
Average						**43,650**

Variance of the Hypothetical Mean Pure Premiums

The computation of the VHM for the pure premiums is similar to that for frequency. One has to first compute the mean pure premium for each type.

For type 1, the mean of the Bernoulli frequency is $p = .4$, and the mean of the Gamma severity is $\alpha/\lambda = 4/.01 = 400$. Thus since frequency and severity are assumed to be independent, the mean pure premium = (Mean Frequency)(Mean Severity) = (.4)(400) = 160. For type 2 , the mean pure premium = (.7)(300) = 210. For type 3, the mean pure premium[29] = (.8)(200) = 160.

One computes the first and second moments of the mean pure premiums as follows:

Class	A Priori Probability	Mean Frequency	Mean Severity	Mean Pure Premium	Square of Pure Premium
1	50%	0.4	400	160	25,600
2	30%	0.7	300	210	44,100
3	20%	0.8	200	160	25,600
Average				175	31,150

Thus the variance of the hypothetical mean pure premiums $= 31{,}150 - 175^2 = 525$.

Estimating the Variance of the Hypothetical Means in the Case of Poisson Frequencies

In real-world applications involving Poisson frequencies it is commonly the case that one estimates the Total Variance

[29]Note that in this example it turns out that the mean pure premium for type 3 happens to equal that for type 1, even though the two types have different mean frequencies and severities. The mean pure premiums tend to be similar when, as in this example, high frequency is associated with low severity.

and the Expected Value of the Process Variance and then estimates the Variance of the Hypothetical Means via: VHM = Total Variance – EPV.

For example, assume that one observes that the claim count distribution is as follows for a large group of insureds:

Total Claim Count:	0	1	2	3	4	5	>5
Percentage of Insureds:	60.0%	24.0%	9.8%	3.9%	1.6%	0.7%	0%

One can estimate the total mean as .652 and the total variance as: $1.414 - .652^2 = .989$.

A	B	C	D
Number of Claims	A Priori Probability	Col. A $\times$ Col. B	Square of Col. A $\times$ Col. B
0	0.60000	0.00000	0.00000
1	0.24000	0.24000	0.24000
2	0.09800	0.19600	0.39200
3	0.03900	0.11700	0.35100
4	0.01600	0.06400	0.25600
5	0.00700	0.03500	0.17500
Sum	1	0.652	1.41400

Assume in addition that the claim count, X, for each individual insured has a Poisson distribution that does not change over time. In other words, each insured's frequency process is given by a Poisson with parameter θ, with θ varying over the group of insureds. Then since the Poisson has its variance equal to its mean, the process variance for each insured is θ; i.e., $\mathrm{Var}[X \mid \theta] = \theta$. Thus the expected value of the process variance is estimated as follows: $E_\theta[\mathrm{Var}[X \mid \theta]] = E_\theta[\theta]$ = overall mean = .652.

Thus we estimate the Variance of the Hypothetical Means as:

$$\text{Total Variance} - \text{EPV} = .989 - .652 = .337.$$

3.1. Exercises

Use the following information for the next two questions:

There are three types of risks. Assume 60% of the risks are of Type A, 25% of the risks are of Type B, and 15% of the risks are of Type C. Each risk has either one or zero claims per year.

Type of Risk	Chance of a Claim	A Priori Chance of Type of Risk
A	20%	60%
B	30%	25%
C	40%	15%

3.1.1. What is the Expected Value of the Process Variance?

3.1.2. What is the Variance of the Hypothetical Means?

Use the following information for the next two questions:

An insured population consists of 9% youthful drivers and 91% adult drivers. Based on experience, we have derived the following probabilities that an individual driver will have n claims in a year's time:

n	Youthful	Adult
0	.80	0.90
1	.15	0.08
2	.04	0.02
3	.01	0.00

3.1.3. What is the Expected Value of the Process Variance?

3.1.4. What is the Variance of the Hypothetical Means?

The following information pertains to the next three questions:

The claim count distribution is as follows for a large sample of insureds:

Total Claim Count	0	1	2	3	4	>4
Percentage of Insureds	55%	30%	10%	4%	1%	0%

Assume the claim count for each individual insured has a Poisson distribution that does not change over time.

3.1.5. What is the Expected Value of the Process Variance?

3.1.6. What is the Total Variance?

3.1.7. What is the Variance of the Hypothetical Means?

3.1.8. The hypothetical mean frequencies of the members of a class of risks are distributed uniformly on the interval (0,10]. The Exponential probability density function for severity, $f(x)$, is defined below, with the r parameter being different for different individuals. r is distributed on (0,2] by the function $g(r)$.

$$f(x) = (1/r)\exp(-x/r) \qquad x \geq 0$$
$$g(r) = r/2 \qquad 0 \leq r \leq 2$$

The frequency and severity are independently distributed. What is the variance of the hypothetical mean pure premiums for this class of risks?

Use the following information for the next six questions:

Two six-sided dice, A_1 and A_2, are used to determine the number of claims. Each side of both dice are marked with either a 0 or a 1, where 0 represents no claim and 1 represents a claim.

The probability of a claim for each die is:

Die	Probability of Claim
A_1	$2/6$
A_2	$3/6$

In addition, there are two spinners, B_1 and B_2, representing claim severity. Each spinner has two areas marked 20 and 50. The probabilities for each claim size are:

	Claim Size	
Spinner	20	50
B_1	.60	.40
B_2	.20	.80

A single observation consists of selecting a die randomly from A_1 and A_2 and a spinner randomly from B_1 and B_2, rolling the selected die, and if there is a claim spinning the selected spinner.

3.1.9. Determine the Expected Value of the Process Variance for the frequency.

3.1.10. Determine the Variance of the Hypothetical Mean frequencies.

3.1.11. Determine the Expected Value of the Process Variance for the severity.

3.1.12. Determine the Variance of the Hypothetical Mean severities.

3.1.13. Determine the Expected Value of the Process Variance for the pure premium.

3.1.14. Determine the Variance of the Hypothetical Mean pure premiums.

Use the following information for the next six questions:

For an individual insured, frequency and severity are independent.

For an individual insured, frequency is given by a Poisson Distribution.

For an individual insured, severity is given by an Exponential Distribution.

Each type is homogeneous; i.e., every insured of a given type has the same frequency process and severity process.

Type	Portion of Insureds in this Type	Mean Frequency	Mean Severity
1	40%	6	100
2	35%	7	125
3	25%	9	200

3.1.15. What is the Expected Value of the Process Variance for the frequency?

3.1.16. What is the Variance of the Hypothetical Mean frequencies?

3.1.17. What is the Expected Value of the Process Variance for the severity?

3.1.18. What is the Variance of the Hypothetical Mean severities?

3.1.19. What is the Expected Value of the Process Variance for the pure premium?

3.1.20. What is the Variance of the Hypothetical Mean pure premiums?

Use the following information for the next two questions:

The probability of y successes in five trials is given by a binomial distribution with parameters 5 and p. The prior distribution of p is uniform on $[0,1]$.

3.1.21. What is the Expected Value of the Process Variance?

3.1.22. What is the Variance of the Hypothetical Means?

3.2. Bühlmann Credibility

Let's continue along with the simple example involving multi-sided dice:

There are a total of 100 multi-sided dice of which 60 are 4-sided, 30 are 6-sided and 10 are 8-sided. The multi-sided dice with 4 sides have 1, 2, 3 and 4 on them. The multi-sided dice with the usual 6 sides have numbers 1 through 6 on them. The multi-sided dice with 8 sides have numbers 1 through 8 on them. For a given die each side has an equal chance of being rolled; i.e., the die is fair.

Your friend has picked at random a multi-sided die. He then rolled the die and told you the result. You are to estimate the result when he *rolls that same die again*.

Using Bühlmann Credibility, the new estimate $= Z(\text{observation}) + (1 - Z)(\text{prior mean})$.

In this example the prior mean is 3. This is the a priori expected value if selecting a die at random and rolling it: $.6(2.5) + .3(3.5) + .1(4.5) = 3.00$. However, since your friend told you additional information about the die he selected, i.e., the value of the first roll, you can come up with a better estimate for the value of the second roll using Bühlmann Credibility.

The prior mean, or a priori expected value, serves as the "other information" to which we apply the complement of credibility. To the extent that our observed data is not credible, we would rely on the prior mean for our estimate. The prior mean reflects what we know about the whole population of dice.

The **Bühlmann Credibility Parameter** is calculated as K = EPV/VHM:

$$K = \textbf{Expected Value of Process Variance/} \\ \textbf{Variance of Hypothetical Means}, \qquad (3.2.1)$$

where the Expected Value of the Process Variance and the Variance of the Hypothetical Means are each calculated for a *single* observation of the risk process.

In this case[30] $K = \text{EPV}/\text{VHM} = 2.15/.45 = 4.778 = 43/9$.

Then for N observations, the Bühlmann Credibility is:

$$Z = N/(N + K) \qquad (3.2.2)$$

In this case for one observation, $Z = 1/(1 + 4.778) = .1731 = 9/52 = .45/(.45 + 2.15)$. Thus in this case if we observe a roll of a 5, then the new estimate is: $(.1731)(5) + (1 - .1731)(3) = 3.3462$. The Bühlmann Credibility estimate is a linear function of the observation: .1731 (observation) + 2.4807.

Observation	1	2	3	4	5	6	7	8
New Estimate	2.6538	2.8269	3	3.1731	3.3462	3.5193	3.6924	3.8655

Note that if $N = 1$, then $Z = 1/(1+K) = \text{VHM}/(\text{VHM}+\text{EPV})$ = VHM/Total Variance.

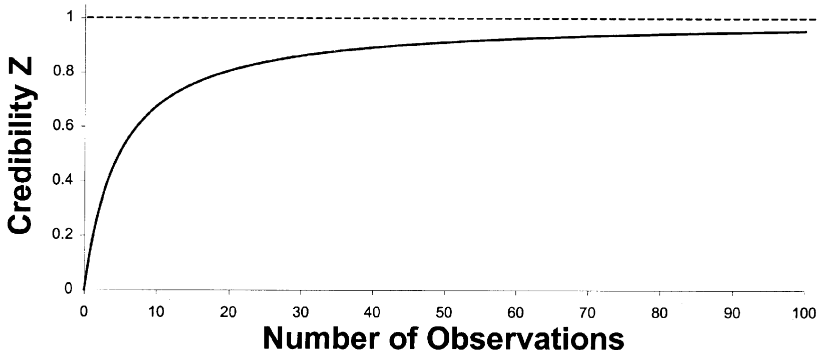
Number of Observations

It makes sense to assign more credibility to more rolls of the selected die, since as we gather more information we should be able to get a better idea of which type of die has been chosen. If one has N observations of the risk process, one assigns Bühlmann Credibility of $Z = N/(N + K)$. For example, with K =

[30]The EPV and the VHM were calculated in Section 3.1.

GRAPH 2

BÜHLMANN CREDIBILITY



4.778, for three observations we have $Z = 3/(3 + 4.778) = .386$. For the Bühlmann Credibility formula as $N \to \infty$, $Z \to 1$, but Bühlmann Credibility never quite reaches 100%. In this example with $K = 4.778$:

Number of Observations	1	2	3	4	5	10	25	100	1,000	10,000
Credibility	17.3%	29.5%	38.6%	45.6%	51.1%	67.7%	84.0%	95.4%	99.5%	99.95%

Graph 2 shows the Bühlmann Credibility. Note that unlike Classical Credibility, Z never reaches 1.00.

If we add up N independent rolls of the same die, the process variances add. So if η^2 is the expected value of the process variance of a single die, then $N\eta^2$ is the expected value of the process variance of the sum of N identical dice. The process variance of one 6-sided die is 35/12, while the process variance of the sum of ten 6-sided dice is 350/12.

In contrast if τ^2 is the variance of the hypothetical means of one die roll, then the variance of the hypothetical means of the sum of N rolls of the *same* die is $N^2\tau^2$. This follows from the fact that each of the means is multiplied by N, and that multiplying by a constant multiplies the variance by the square of that constant.

Thus as N increases, the variance of the hypothetical means of the sum goes up as N^2 while the process variance goes up only as N. Based on the case with one roll, we expect the credibility to be given by Z = VHM/Total Variance = VHM/(VHM + EPV) = $N^2\tau^2/(N^2\tau^2 + N\eta^2) = N/(N + \eta^2/\tau^2) = N/(N + K)$, where $K = \eta^2/\tau^2$ = EPV/VHM, with EPV and VHM each for a *single* die.

In general **one computes the EPV and VHM for a single observation and then plugs into the formula for Bühlmann Credibility the number of observations N. If one is estimating claim frequencies or pure premiums then N is in exposures. If one is estimating claim severities then N is in number of claims. (N is in the units of whatever is in the denominator of the quantity one is estimating.**[31])

A Series of Examples

In section 3.1, the following information was used in a series of examples involving the frequency, severity, and pure premium:

Type	Portion of Risks in this Type	Bernoulli (Annual) Frequency Distribution	Gamma Severity Distribution
1	50%	$p = 40\%$	$\alpha = 4,\ \lambda = .01$
2	30%	$p = 70\%$	$\alpha = 3,\ \lambda = .01$
3	20%	$p = 80\%$	$\alpha = 2,\ \lambda = .01$

We assume that the types are homogeneous; i.e., every insured of a given type has the same frequency and severity process. Assume that for an individual insured, frequency and severity are independent.

[31] Claim frequency = claims/exposures. Claim severity = losses/claims. Pure Premium = loss/exposures.

Using the Expected Value of the Process Variance and the Variance of the Hypothetical Means computed in the previous section, one can compute the Bühlmann Credibility Parameter in each case.

Suppose that an insured is picked at random and we do not know what type he is.[32] For this randomly selected insured during 4 years one observes 3 claims for a total[33] of \$450. Then one can use Bühlmann Credibility to predict the future frequency, severity, or pure premium of this insured.

Frequency Example

As computed in section 3.1, the EPV of the frequency = .215, while the variance of the hypothetical mean frequencies = .0301. Thus the Bühlmann Credibility parameter is: K = EPV/VHM = .215/.0301 = 7.14.

Thus 4 years of experience are given a credibility of $4/(4+K) = 4/11.14 = 35.9\%$. The observed frequency is $3/4 = .75$. The a priori mean frequency is .57. Thus the estimate of the future frequency for this insured is $(.359)(.75) + (1 - .359)(.57) = .635$.

Severity Example

As computed in section 3.1, the EPV of the severity = 30,702, while the variance of the hypothetical mean severities = 6,265. Thus the Bühlmann Credibility parameter is K = EPV/VHM = 30,702/6,265 = 4.90.

Thus 3 observed claims[34] are given a credibility of $3/(3+K) = 3/7.9 = 38.0\%$. The observed severity is \$450/3 = \$150. The

[32]The latter is very important. If one knew which type the insured was, one would use the expected value for that type to estimate the future frequency, severity, or pure premium.
[33]Unlike the Bayesian Analysis case to be covered subsequently, even if one were given the separate claim amounts, the Bühlmann Credibility estimate of severity only makes use of the sum of the claim amounts.
[34]Note that the number of observed *claims* is used to determine the Bühlmann Credibility of the severity.

a priori mean severity is \$307. Thus the estimate of the future severity for this insured is $(.380)(150) + (1 - .380)(307) =$ \$247.3.

Pure Premium Example

As computed in section 3.1, the EPV of the pure premium = 43,650, while the variance of the hypothetical mean pure premiums = 525. Thus the Bühlmann Credibility parameter is $K = \text{EPV}/\text{VHM} = 43{,}650/525 = 83.1$.

Thus 4 years of experience are given a credibility of $4/(4 + K) = 4/87.1 = 4.6\%$. The observed pure premium is \$450/4 = \$112.5. The a priori mean pure premium is \$175. Thus the estimate of the future pure premium for this insured is: $(.046)(112.5) + (1 - .046)(175) =$ \$172.

Note that this estimate of the future pure premium is *not* equal to the product of our previous estimates of the future frequency and severity. (.635)(\$247.3) = \$157 $\neq$ \$172. In general one does not get the same result if one uses credibility to make separate estimates of the frequency and severity instead of directly estimating the pure premium.

Assumptions Underlying $Z = N/(N + K)$

There are a number of important assumptions underlying the formula $Z = N/(N + K)$ where $K = \text{EPV}/\text{VHM}$. While these assumptions generally hold in this chapter, they hold in many, but far from every, real-world application.[35] These assumptions are:

1. The complement of credibility is given to the overall mean.
2. The credibility is determined as the slope of the weighted least squares line to the Bayesian Estimates.

[35] See for example, Howard Mahler's discussion of Glenn Meyers' "An Analysis of Experience Rating," *PCAS*, 1987.

3. The risk parameters and risk process do not shift over time.[36]

4. The expected value of the process variance of the sum of N observations increases as N. Therefore the expected value of the process variance of the average of N observations decreases as $1/N$.

5. The variance of the hypothetical means of the sum of N observations increases as N^2. Therefore the variance of the hypothetical means of the average of N observations is independent of N.

In addition one must be careful that an insured has been picked at random, that we observe that insured and then we attempt to make an estimate of the future observation of that *same* insured. If instead one goes back and chooses a new insured at random, then the information contained in the observation has been lost.

3.2. Exercises

3.2.1. The Expected Value of the Process Variance is 100. The Variance of the Hypothetical Means is 8. How much Bühlmann Credibility is assigned to 20 observations of this risk process?

3.2.2. If 5 observations are assigned 70% Bühlmann Credibility, what is the value of the Bühlmann Credibility parameter K?

3.2.3. Your friend picked at random one of three multi-sided dice. He then rolled the die and told you the result. You are to estimate the result when he rolls that same die again. One of the three multi-sided dice has 4 sides (with

[36] In the Philbrick Target Shooting Example discussed in a subsequent section, we assume the targets are fixed and that the skill of the marksmen does not change over time.

1, 2, 3 and 4 on them), the second die has the usual 6 sides (numbered 1 through 6), and the last die has 8 sides (with numbers 1 through 8). For a given die each side has an equal chance of being rolled, i.e. the die is fair. Assume the first roll was a five. Use Bühlmann Credibility to estimate the next roll of the same die.

Hint: The mean of a die with S sides is: $(S+1)/2$. The variance of a die with S sides is: $(S^2-1)/12$.

Use the following information for the next two questions:

There are three large urns, each filled with so many balls that you can treat it as if there are an infinite number. Urn 1 contains balls with "zero" written on them. Urn 2 has balls with "one" written on them. The final Urn 3 is filled with 50% balls with "zero" and 50% balls with "one." An urn is chosen at random and five balls are selected.

3.2.4. If all five balls have "zero" written on them, use Bühlmann Credibility to estimate the expected value of another ball picked from that urn.

3.2.5. If three balls have "zero" written on them and two balls have "one" written on them, use Bühlmann Credibility to estimate the expected value of another ball picked from that urn.

3.2.6. There are two types of urns, each with many balls labeled \$1,000 and \$2,000.

Type of Urn	A Priori chance of This Type of Urn	Percentage of \$1,000 Balls	Percentage of \$2,000 Balls
I	80%	90%	10%
II	20%	70%	30%

An urn is selected at random, and you observe a total of $8,000 on 5 balls drawn from that urn at random. Using Bühlmann Credibility, what is estimated value of the next ball drawn from that urn?

3.2.7. The aggregate loss distributions for three risks for one exposure period are as follows:

	Aggregate Losses		
Risk	$0	$100	$500
A	0.90	0.07	0.03
B	0.50	0.30	0.20
C	0.30	0.33	0.37

A risk is selected at random and is observed to have $500 of aggregate losses in the first exposure period. Determine the Bühlmann Credibility estimate of the expected value of the aggregate losses for the same risk's second exposure period.

3.2.8. A die is selected at random from an urn that contains four 6-sided dice with the following characteristics:

Number	Number of Faces			
on Face	Die A	Die B	Die C	Die D
1	3	1	1	1
2	1	3	1	1
3	1	1	3	1
4	1	1	1	3

The first five rolls of the selected die yielded the following in sequential order: 2, 3, 1, 2, and 4. Using Bühlmann Credibility, what is the expected value of the next roll of the same die?

Use the following information for the following seven questions:

There are three types of drivers with the following characteristics:

Type	Portion of Drivers of This Type	Poisson Annual Claim Frequency	Pareto Claim Severity*
Good	60%	5%	$\alpha = 5,\ \lambda = 10{,}000$
Bad	30%	10%	$\alpha = 4,\ \lambda = 10{,}000$
Ugly	10%	20%	$\alpha = 3,\ \lambda = 10{,}000$

For any individual driver, frequency and severity are independent.

3.2.9. A driver is observed to have over a five-year period a single claim. Use Bühlmann Credibility to predict this driver's future annual claim frequency.

3.2.10. What is the expected value of the process variance of the claim severities (for the observation of a single claim)?

3.2.11. What is the variance of the hypothetical mean severities (for the observation of a single claim)?

3.2.12. Over several years, for an individual driver you observe a single claim of size $25,000. Use Bühlmann Credibility to estimate this driver's future average claim severity.

3.2.13. What is the expected value of the process variance of the pure premiums (for the observation of a single exposure)?

3.2.14. What is the variance of the hypothetical mean pure premiums (for the observation of a single exposure)?

3.2.15. A driver is observed to have over a five-year period a total of $25,000 in losses. Use Bühlmann Credibility to predict this driver's future pure premium.

*See the Appendix on claim frequency and severity distributions.

3.2.16. There are two classes of insureds in a given population. Each insured has either no claims or exactly one claim in one experience period. For each insured the distribution of the number of claims is binomial. The probability of a claim in one experience period is 0.20 for Class 1 insureds and 0.30 for Class 2. The population consists of 40% Class 1 insureds and 60% for Class 2. An insured is selected at random without knowing the insured's class. What credibility would be given to this insured's experience for five experience periods using Bühlmann's Credibility Formula?

Use the following information for the next two questions:

	Number of Claims		Size of Loss	
Class	Mean	Process Variance	Mean	Variance
A	.1667	.1389	4	20
B	.8333	.1389	2	5

Each class is homogeneous with all members of the class having the same mean and process variance. Frequency and severity are independently distributed. Classes A and B have the same number of risks. A risk is randomly selected from one of the two classes and four observations are made of the risk.

3.2.17. Determine the value for the Bühlmann Credibility, Z, that can be applied to the observed pure premium.

3.2.18. The pure premium calculated from the four observations is 0.25. Determine the Bühlmann Credibility estimate for the risk's pure premium.

3.2.19. You are given the following:

- X is a random variable with mean m and variance v.
- m is a random variable with mean 2 and variance 4.

- v is a random variable with mean 8 and variance 32.

Determine the value of the Bühlmann Credibility factor Z, after three observations of X. (m and v are constant during the observation periods.)

3.3. Target Shooting Example

In the classic paper by Stephen Philbrick[37] there is an excellent target shooting example that illustrates the ideas of Bühlmann Credibility. Assume there are four marksmen each shooting at his own target. Each marksman's shots are assumed to be distributed around his target, marked by one of the letters A, B, C, and D, with an expected mean equal to the location of his target. Each marksman is shooting at a different target.

If the targets are arranged as in Figure 1, the resulting shots of each marksman would tend to cluster around his own target. The shots of each marksman have been distinguished by a different symbol. So for example the shots of marksman B are shown as triangles. We see that in some cases one would have a hard time deciding which marksman had made a particular shot if we did not have the convenient labels.

The point E represents the average of the four targets A, B, C, and D. Thus E is the grand mean.[38] If we did not know which marksman was shooting we would estimate that the shot would be at E; the a priori estimate is E.

Once we observe a shot from an unknown marksman,[39] we could be asked to estimate the location of the next shot from

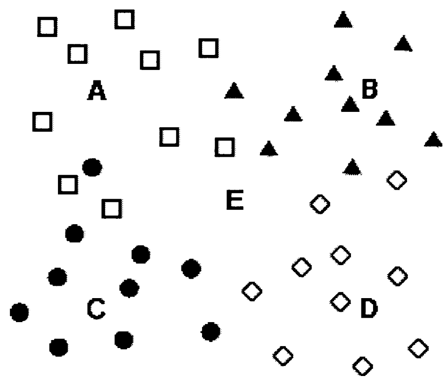
---

[37]"An Examination of Credibility Concepts," *PCAS*, 1981.

[38]In this example, each of the marksmen is equally likely; that is, they fire the same number of shots. Thus we weight each target equally. As was seen previously, in general one would take a weighted average using the not necessarily equal a priori probabilities as the weights.

[39]Thus the shot does *not* have one of the convenient labels attached to it. This is analogous to the situation in Auto Insurance, where the drivers in a classification are presumed not to be wearing little labels telling us who are the safer and less safe drivers in the class. We rely on the observed experience to help estimate that.

FIGURE 1

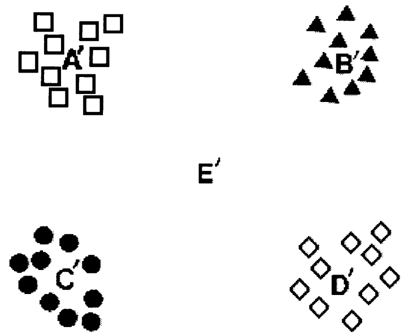


the *same* marksman. Using Bühlmann Credibility our estimate would be between the observation and the a priori mean of E. The larger the credibility assigned to the observation, the closer the estimate is to the observation. The smaller the credibility assigned to the data, the closer the estimate is to E.

There are a number of features of this target shooting example that control how much Bühlmann Credibility is assigned to our observation. We have assumed that the marksmen are not perfect; they do not always hit their target. The amount of spread of their shots around their targets can be measured by the variance. The average spread over the marksmen is the Expected Value of the Process Variance (EPV). The better the marksmen, the smaller the EPV and the more tightly clustered around the targets the shots will be.

The worse the marksmen, the larger the EPV and the less tightly the shots are spread. The better the marksmen, the more information contained in a shot. The worse the marksmen, the more random noise contained in the observation of the location of a shot. Thus when the marksmen are good, we expect to give more weight to an observation (all other things being equal) than

FIGURE 2



when the marksmen are bad. Thus the better the marksmen, the higher the credibility:

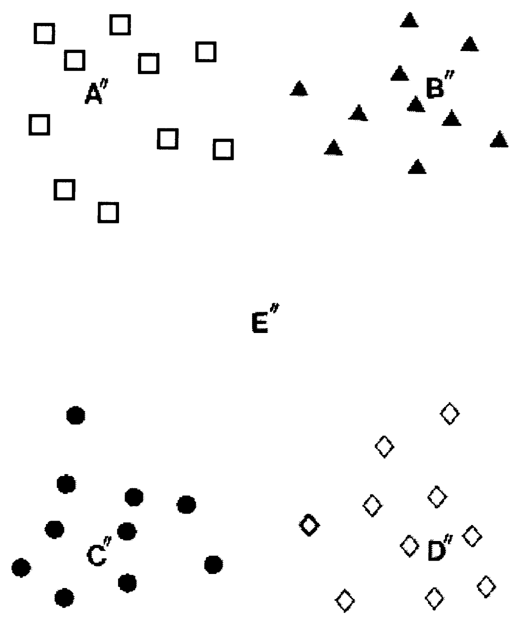
Marksmen	Clustering of Shots	Expected Value of the Process Variance	Information Content	Credibility Assigned to an Observation
Good	Tight	Small	High	Larger
Bad	Loose	Large	Low	Smaller

The smaller the Expected Value of the Process Variance the larger the credibility. This is illustrated by Figure 2. It is assumed in Figure 2 that each marksman is better[40] than was the case in Figure 1. The EPV is smaller and we assign more credibility to the observation. This makes sense, since in Figure 2 it is a lot easier to tell which marksman is likely to have made a particular shot based solely on its location.

Another feature that determines how much credibility to give an observation is how far apart the four targets are placed. As we move the targets further apart (all other things being equal) it is easier to distinguish the shots of the different marksmen. Each

[40] Alternatively, the marksmen could be shooting from closer to the targets.

FIGURE 3



target is a hypothetical mean of one of the marksmen's shots. The spread of the targets can be quantified as the Variance of the Hypothetical Means.

Targets	Variance of the Hypothetical Means	Information Content	Credibility Assigned to an Observation
Closer	Small	Low	Smaller
Further Apart	Large	High	Larger

As illustrated in Figure 3, the further apart the targets the more credibility we would assign to our observation. The larger the VHM the larger the credibility. It is easier to distinguish which marksman made a shot based solely on its location in Figure 3 than in Figure 1.

The third feature that one can vary is the number of shots observed from the *same* unknown marksman. The more shots we observe, the more information we have and thus the more credibility we would assign to the average of the observations.

Each of the three features discussed is reflected in the formula for Bühlmann Credibility $Z = N/(N + K) = N(\text{VHM})/\{N(\text{VHM}) + \text{EPV}\}$. Thus, as the EPV increases, Z decreases; as VHM increases, Z increases; and as N increases, Z increases.

Feature of Target Shooting Example	Mathematical Quantification	Bühlmann Credibility
Better Marksmen	Smaller EPV	Larger
Targets Further Apart	Larger VHM	Larger
More Shots	Larger N	Larger

Expected Value of the Process Variance vs. Variance of the Hypothetical Means

There are two separate reasons why the observed shots vary. First, the marksmen are not perfect. In other words the Expected Value of the Process Variance is positive. Even if all the targets were in the same place, there would still be a variance in the observed results. This component of the total variance due to the imperfection of the marksmen is quantified by the EPV.

Second, the targets are spread apart. In other words, the Variance of the Hypothetical Means is positive. Even if every marksman were perfect, there would still be a variance in the observed results, when the marksmen shoot at different targets. This component of the total variance due to the spread of the targets is quantified by the VHM.

One needs to understand the distinction between these two sources of variance in the observed results. Also one has to know that the total variance of the observed shots is a sum of these two components: Total Variance = EPV + VHM.

Bühlmann Credibility is a Relative Concept

In Philbrick's target shooting example, we are comparing two estimates of the location of the next shot from the same marksman. One estimate is the average of the observations; the other estimate is the average of the targets, the a priori mean.

When the marksmen are better, there is less random fluctuation in the shots and the average of the observations is a better estimate, relative to the a priori mean. In this case, the weight Z applied to the average of the observations is larger, while $(1 - Z)$ applied to the a priori mean is smaller.

As the targets get closer together, there is less variation of the hypothetical means, and the a priori mean becomes a better estimate, relative to the average of the observations. In this case, the weight Z applied to the average of the observations is smaller, while the complement $(1 - Z)$ applied to the a priori mean is larger.

Bühlmann Credibility measures the usefulness of one estimator, the average of the observations, relative to another estimator, the a priori mean. If $Z = 50\%$, then the two estimators are equally good or equally bad. In general, Bühlmann Credibility is a relative measure of the value of the information contained in the observation versus that in the a priori mean.

A One-Dimensional Target Shooting Example

Assume a one-dimensional example of Philbrick's target shooting model such that the marksmen only miss to the left or right. Assume:

- There are four marksmen.
- The targets for the marksmen are at the points on the number line: 10, 20, 30, and 40.
- The accuracy of each marksman follows a normal distribution with mean equal to his target value and with standard deviation of 12.

Assume a single shot at 18. To use Bühlmann Credibility we need to calculate the Expected Value of the Process Variance and the Variance of the Hypothetical Means. The process variance for every marksman is assumed to be the same and equal to $12^2 = 144$. Thus the EPV = 144. The overall mean is 25 and the variance is $(1/4)(15^2 + 5^2 + 5^2 + 15^2) = 125$. Thus the Bühlmann Credibility parameter is $K = \mathrm{EPV}/\mathrm{VHM} = 144/125 = 1.152$. The credibility of a single observation is $Z = 1/(1 + 1.152) = 46.5\%$. Thus if one observes a single shot at 18, then the Bühlmann Credibility estimate of the next shot is $(18)(46.5\%) + (25)(53.5\%) = 21.7$.

More Shots

What if instead of a single shot at 18 one observed three shots at 18, 26 and 4 from the same unknown marksman?

As calculated above $K = 1.152$. The credibility of 3 observations is $Z = 3/(3 + 1.152) = 72.3\%$. The larger number of observations has increased the credibility. The average of the observations is $(18 + 26 + 4)/3 = 16$. The a priori mean is 25. Thus the Bühlmann Credibility estimate of the next shot is: $(16)(72.3\%) + (25)(27.7\%) = 18.5$.

Moving the Targets

Assume that the targets were further apart. Assume:

- There are four marksmen.
- The targets for the marksmen are at the points on the number line: 20, 40, 60, and 80.
- The accuracy of each marksman follows a normal distribution with mean equal to his target value and with standard deviation of 12.

Then each shot has more informational content about which marksman produced it. Assume we observe three shots from an unknown marksman at: 38, 46 and 24. The EPV is still 144 while the VHM is now $(4)(125) = 500$, so the Bühlmann Credibility

Parameter $K = \text{EPV}/\text{VHM} = 144/500 = .288$. Thus the credibility assigned to 3 shots is $3/3.288 = 91.2\%$, larger than before. The larger VHM has increased the credibility.[41] The average of the observations is $(38 + 46 + 24)/3 = 36$. The a priori mean is 50. Thus the Bühlmann Credibility estimate of the next shot is: $(36)(91.2\%) + (50)(8.8\%) = 37.2$.

Altering the Skill of the Marksmen

Assume that the marksmen are more skilled.[42] Assume:

- There are four marksmen.
- The targets for the marksmen are at the points on the number line: 10, 20, 30, and 40.
- The accuracy of each marksman follows a normal distribution with mean equal to his target value and with standard deviation of 6.

With a smaller process variance, each shot has more informational content about which marksman produced it. Assume we observe three shots from an unknown marksman at: 18, 26 and 4. The EPV is $6^2 = 36$ while the VHM is 125, so the Bühlmann Credibility $K = \text{EPV}/\text{VHM} = 36/125 = .288$. Thus the credibility assigned to 3 shots is $3/3.288 = 91.2\%$, more than in the original example. The smaller EPV has increased the credibility.[43] The average of the observations is $(18 + 26 + 4)/3 = 16$. The a priori mean is 25. Thus the Bühlmann Credibility estimate of the next shot is: $(16)(91.2\%) + (25)(8.8\%) = 16.8$.

Limiting Situations and Bühlmann Credibility

As the number of observations approaches infinity, the credibility approaches one. In the target shooting example, as the

[41] If instead one had moved the targets closer together, then the credibility assigned to a single shot would have been less. A smaller VHM leads to less credibility.

[42] Alternatively, assume the marksmen are shooting from closer to the targets.

[43] If instead one had less skilled marksmen, then the credibility assigned to a single shot would have been less. A larger EPV leads to less credibility.

number of shots approaches infinity, our Bühlmann Credibility estimate approaches the mean of the observations.

On the other hand, if we have no observations, then the estimate is the a priori mean. We give the a priori mean a weight of 1, so $1 - Z = 1$ or $Z = 0$.

Bühlmann Credibility is given by $Z = N/(N + K)$. In the usual situations where one has a finite number of observations, $0 < N < \infty$, one will have $0 < Z < 1$ provided $0 < K < \infty$. The Bühlmann Credibility is only zero or unity in unusual situations.

The Bühlmann Credibility parameter $K = \text{EPV}/\text{VHM}$. So $K = 0$ if $\text{EPV} = 0$ or $\text{VHM} = \infty$. On the other hand K is infinite if $\text{EPV} = \infty$ or $\text{VHM} = 0$.

The Expected Value of the Process Variance is zero only if one has certainty of results.[44] In the case of the Philbrick Target Shooting Example, if all the marksmen were absolutely perfect, then the expected value of the process variance would be zero. In that situation we assign the observation a credibility of unity; our new estimate is the observation.

The Variance of the Hypothetical Means is infinite if one has little or no knowledge and therefore has a large variation in hypotheses.[45] In the case of the Philbrick Target Shooting Example, as the targets get further and further apart, the variance of the hypothetical means approaches infinity. We assign the observations more and more weight as the targets get further apart. If one target were in Alaska, another in California, another in Maine and the fourth in Florida, we would give the observation virtually 100% credibility. In the limit, our new estimate is the observation; the credibility is one.

---

[44] For example, one could assume that it is certain that the sun will rise tomorrow; there has been no variation of results, the sun has risen every day of which you are aware.

[45] For example, an ancient Greek philosopher might have hypothesized that the universe was more than 3,000 years old with all such ages equally likely.

However, in most applications of Bühlmann Credibility the Expected Value of the Process Variance is positive and the Variance of the Hypothetical Means is finite, so that $K > 0$.

The Expected Value of the Process Variance can be infinite only if the process variance is infinite for at least one of the types of risks. If in an example involving claim severity, one assumed a Pareto distribution with $\alpha \leq 2$, then one would have infinite process variance. In the Philbrick Target Shooting Example, a marksman would have to be infinitely terrible in order to have an infinite process variance. As the marksmen get worse and worse, we give the observation less and less weight. In the limit where the location of the shot is independent of the location of the target we give the observation no weight; the credibility is zero.

The Variance of the Hypothetical Means is zero only if all the types of risks have the same mean. For example, in the Philbrick Target Shooting Example, if all the targets are at the same location (or alternatively each of the marksmen is shooting at the same target) then the VHM = 0. As the targets get closer and closer to each other, we give the observation less and less weight. In the limit, we give the observation no weight; the credibility is zero. In the limit, all the weight is given to the single target.

However, in the usual applications of Bühlmann Credibility there is variation in the hypotheses, and there is a finite expected value of process variance, and therefore K is finite.

Assuming $0 < K < \infty$ and $0 < N < \infty$, then $0 < Z < 1$. **Thus in ordinary circumstances the Bühlmann Credibility is strictly between zero and one.**

3.3. Exercises

3.3.1. In which of the following should credibility be expected to increase?

1. Larger quantity of observations.

2. Increase in the prior mean.
3. Increase in the variance of hypothetical means.

3.3.2. There are three marksmen, each of whose shots are Normally Distributed (in one dimension) with means and standard deviations:

Risk	Mean	Standard Deviation
A	10	3
B	20	5
C	30	15

A marksman is chosen at random. You observe two shots at 10 and 14. Using Bühlmann Credibility estimate the next shot from the same marksman.

Use the following information to answer the next two questions:

Assume you have two shooters, each of whose shots is given by a (one dimensional) Normal distribution:

Shooter	Mean	Variance
A	+1	1
B	−1	25

Assume a priori that each shooter is equally likely

3.3.3. You observe a single shot at +4. Use Bühlmann Credibility to estimate the location of the next shot.

3.3.4. You observed three shots at 2, 0, 1. Use Bühlmann Credibility to estimate the location of the next shot.

4. BAYESIAN ANALYSIS

Bayesian Analysis is another technique to update a prior hypothesis based on observations, closely related to the use of

Bühlmann Credibility. In fact, the use of Bühlmann Credibility is the least squares linear approximation to Bayesian Analysis.[46] First, some preliminary mathematical ideas related to Bayesian Analysis will be discussed.

4.1. Mathematical Preliminaries

Conditional Distributions

Example 4.1.1: Assume that 14% of actuarial students take exam seminars and that 8% of actuarial students both take exam seminars and pass their exam. What is the chance of a student who has taken an exam seminar passing his/her exam?

[*Solution:* $8\%/14\% = 57\%$. (Assume 1,000 total students, of whom 140 take exam seminars. Of these 140 students, 80 pass, for a pass rate of 57%.)]

This is a simple example of a conditional probability.

The conditional probability of an event A given another event B is defined as:

$$P[A \mid B] = P[A \text{ and } B]/P[B] \tag{4.1.1}$$

In the simple example, $A = \{\text{student passes exam}\}$, $B = \{\text{student takes exam seminar}\}$, $P[A \text{ and } B] = 8\%$, $P[B] = 14\%$. Thus $P[A|B] = P[A \text{ and } B]/P[B] = 8\%/14\% = 57\%$.

Theorem of Total Probability

Example 4.1.2: Assume that the numbers of students taking an exam by exam center are as follows: Chicago 3,500, Los Angeles 2,000, New York 4,500. The number of students from each exam center passing the exam are: Chicago 2,625, Los Angeles 1,200, New York 3,060. What is the overall passing percentage?

[*Solution:* $(2{,}625 + 1{,}200 + 3{,}060)/(3{,}500 + 2{,}000 + 4{,}500) = 6{,}885/10{,}000 = 68.85\%$.]

[46] A proof can be found in Bühlmann, "Experience Rating and Credibility" or Klugman, et al., *Loss Models: From Data to Decisions*.

If one has a set of mutually disjoint events A_i, then one can write the marginal distribution function $P[B]$ in terms of the conditional distributions $P[B \mid A_i]$ and the probabilities $P[A_i]$:

$$P[B] = \sum_i P[B \mid A_i] P[A_i] \tag{4.1.2}$$

This theorem follows from $\sum_i P[B \mid A_i]P[A_i] = \sum_i P[B \text{ and } A_i] = P[B]$, provided that the A_i are disjoint events that cover all possibilities.

Thus one can compute probabilities of events either directly or by summing a product of terms.

Example 4.1.3: Assume that the percentages of students taking an exam by exam center are as follows: Chicago 35%, Los Angeles 20%, New York 45%. The percentages of students from each exam center passing the exam are: Chicago 75%, Los Angeles 60%, New York 68%. What is the overall passing percentage?

[*Solution:* $\sum_i P[B \mid A_i]P[A_i] = (75\%)(35\%) + (60\%)(20\%) + (68\%)(45\%) = 68.85\%$.]

Note that example 4.1.3 is mathematically the same as example 4.1.2. This is a concrete example of the Theorem of Total Probability, Equation 4.1.2.

Conditional Expectation

In general, in order to compute a conditional expectation, we take the weighted average over all the possibilities x:

$$E[X \mid B] = \sum_x x P[X = x \mid B] \tag{4.1.3}$$

Example 4.1.4: Let G be the result of rolling a green 6-sided die. Let R be the result of rolling a red 6-sided die. G and R are independent of each other. Let M be the maximum of G and R. What is the expectation of the conditional distribution of M if $G = 3$?

[*Solution:* The conditional distribution of M if $G = 3$ is: $f(3) = 3/6$, $f(4) = 1/6$, $f(5) = 1/6$, and $f(6) = 1/6$. (Note that if $G = 3$, then $M = 3$ if $R = 1$, 2, or 3. So, $f(3) = 3/6$.) Thus the mean of the conditional distribution of M if $G = 3$ is: $(3)(3/6) + (4)(1/6) + (5)(1/6) + (6)(1/6) = 4$.]

4.1. Exercises

4.1.1. Assume that 5% of men are colorblind, while .25% of women are colorblind. A colorblind person is picked out of a population made up 10% of men and 90% of women. What is the chance the colorblind person is a man?

Use the following information for the next six questions:

A large set of urns contain many black and red balls. There are four types of urns each with differing percentages of black balls. Each type of urn has a differing chance of being picked.

Type of Urn	A Priori Probability	Percentage of Black Balls
I	40%	5%
II	30%	8%
III	20%	13%
IV	10%	18%

4.1.2. An urn is picked and a ball is selected from that urn. What is the chance that the ball is black?

4.1.3. An urn is picked and a ball is selected from that urn. If the ball is black, what is the chance that Urn I was picked?

4.1.4. An urn is picked and a ball is selected from that urn. If the ball is black, what is the chance that Urn II was picked?

4.1.5. An urn is picked and a ball is selected from that urn. If the ball is black, what is the chance that Urn III was picked?

4.1.6. An urn is picked and a ball is selected from that urn. If the ball is black, what is the chance that Urn IV was picked?

4.1.7. An urn is picked and a ball is selected from that urn. If the ball is black, what is the chance that the next ball picked from that same urn will be black?

Use the following information for the next two questions:

V and X are each given by the result of rolling a 6-sided die. V and X are independent of each other. $Y = V + X$.

4.1.8. What is the probability that $X = 5$ if $Y \geq 9$?

4.1.9. What is the expected value of X if $Y \geq 9$?

Use the following information for the next two questions:

City	Percentage of Total Drivers	Percent of Drivers Accident-Free
Boston	40%	80%
Springfield	25%	85%
Worcester	20%	90%
Pittsfield	15%	95%

4.1.10. A driver is picked at random. If the driver is accident-free, what is the chance the driver is from Boston?

4.1.11. A driver is picked at random. If the driver has had an accident, what is the chance the driver is from Pittsfield?

4.1.12. On a multiple choice exam, each question has 5 possible answers, exactly one of which is correct. On those questions for which he is not certain of the answer, Stu

Dent's strategy for taking the exam is to answer at random from the 5 possible answers. Assume he correctly answers the questions for which he knows the answers. If Stu Dent knows the answers to 62% of the questions, what is the probability that he knew the answer to a question he answered correctly?

4.2. Bayesian Analysis

Take the following simple example. Assume there are two types of risks, each with Bernoulli claim frequencies. One type of risk has a 30% chance of a claim (and a 70% chance for no claims.) The second type of risk has a 50% chance of having a claim. Of the universe of risks, 3/4 are of the first type with a 30% chance of a claim, while 1/4 are of the second type with a 50% chance of having a claim.

Type of Risk	A Priori Probability that a Risk is of this Type	Chance of a Claim Occurring for a Risk of this Type
1	3/4	30%
2	1/4	50%

If a risk is chosen at random, then the chance of having a claim is $(3/4)(30\%) + (1/4)(50\%) = 35\%$. In this simple example, there are two possible outcomes: either we observe 1 claim or no claims. Thus the chance of no claims is 65%.

Assume we pick a risk at random and observe no claim. Then what is the chance that we have risk Type 1? By the definition of the conditional probability we have: $P(\text{Type} = 1 \mid n = 0) = P(\text{Type} = 1 \text{ and } n = 0)/P(n = 0)$. However, $P(\text{Type} = 1 \text{ and } n = 0) = P(n = 0 \mid \text{Type} = 1)P(\text{Type} = 1) = (.7)(.75)$. Therefore, $P(\text{Type} = 1 \mid n = 0) = P(n = 0 \mid \text{Type} = 1)P(\text{Type} = 1)/P(n = 0) = (.7)(.75)/.65 = .8077$.

This is a special case of **Bayes' Theorem:**

$$P(A \mid B) = P(B \mid A)P(A)/P(B)$$

$$P(\textbf{Risk Type} \mid \textbf{Obser.}) = P(\textbf{Obser.} \mid \textbf{Risk Type}) \times P(\textbf{Risk Type})/P(\textbf{Obser.}) \qquad (4.2.1)$$

Example 4.2.1: Assume we pick a risk at random and observe no claim. Then what is the chance that we have risk Type 2?

[*Solution:* $P(\text{Type} = 2 \mid n = 0) = P(n = 0 \mid \text{Type} = 2)P(\text{Type} = 2) / P(n = 0) = (.5)(.25)/.65 = .1923$.]

Of course with only two types of risks the chance of a risk being Type 2 is unity minus the chance of being Type 1. The posterior probability that the selected risk is Type 1 is .8077 and the posterior probability that it is Type 2 is .1923.

Posterior Estimates

Now not only do we have probabilities posterior to an observation, but we can use these to estimate the chance of a claim if the same risk is observed again. For example, if we observe no claim the estimated claim frequency for the same risk is: (posterior prob. Type 1)(claim freq. Type 1) + (posterior prob. Type 2)(claim freq. Type 2) = (.8077)(30%) + (.1923)(50%) = 33.85%.

Note that the posterior estimate is a weighted average of the hypothetical means for the different types of risks. Thus the posterior estimate of 33.85% is in the range of the hypotheses, 30% to 50%. This is true in general for Bayesian analysis.

The result of Bayesian Analysis is always within the range of hypotheses.

This is not necessarily true for the results of applying Credibility.

Example 4.2.2: What if a risk is chosen at random and one claim is observed. What is the posterior estimate of the chance of a claim from this same risk?

[*Solution:* (.6429)(.3) + (.3571)(.5) = 37.14%

A	B	C	D	E	F
Type of Risk	A Priori Chance of This Type of Risk	Chance of the Observation	Prob. Weight = Product of Columns B & C	Posterior Chance of This Type of Risk	Mean Annual Freq.
1	0.75	0.3	0.225	64.29%[47]	0.30
2	0.25	0.5	0.125	35.71%	0.50
Overall			0.350	1.000	**37.14%**

For example, $P(\text{Type} = 1 \mid n = 1) = P(\text{Type} = 1 \text{ and } n = 1)/P(n = 1) = (.75)(.3)/.35 = .643$, $P(\text{Type} = 2 \mid n = 1) = P(\text{Type} = 2 \text{ and } n = 1)/P(n = 1) = (.25)(.5)/.35 = .357$.]

Note how the estimate posterior to the observation of one claim is 37.14%, greater than the a priori estimate of 35%. The observation has let us infer that it is more likely that the risk is of the high frequency type than it was prior to the observation. Thus we infer that the future chance of a claim from this risk is higher than it was prior to the observation. Similarly, the estimate posterior to the observation of no claim is 33.85%, less than the a priori estimate of 35%.

We had a 65% chance of observing no claim and a 35% chance of observing a claim. Weighting together the two posterior estimates: (65%)(33.85%) + (35%)(37.14%) = 35%. The weighted average of the posterior estimates is equal to the overall a priori mean. This is referred to as "the estimates being in balance." If D_i are the possible outcomes, then the Bayesian estimates are $E[X \mid D_i]$. Then $\sum_i P(D_i)E[X \mid D_i] = E[X]$ = the a priori mean.

$$\sum_i P(D_i)E[X \mid D_i] = E[X] \tag{4.2.2}$$

[47] Note that .6429 = .225/.350 and .3571 = .125/.350. The a priori chance of the observation is .350. Thus the values in column E are the resluts of applying Bayes Theorem, equation 4.2.1

The estimates that result from Bayesian Analysis are always in balance:

The sum of the product of the a priori chance of each outcome times its posterior Bayesian estimate is equal to the a priori mean.

Multi-Sided Dice Example

Let's illustrate Bayesian Analysis with a simple example involving multi-sided dice:

Assume that there are a total of 100 multi-sided dice of which 60 are 4-sided, 30 are 6-sided and 10 are 8-sided. The multi-sided dice with 4 sides have 1, 2, 3 and 4 on them. The multi-sided dice with the usual 6 sides have numbers 1 through 6 on them. The multi-sided dice with 8 sides have numbers 1 through 8 on them. For a given die each side has an equal chance of being rolled; i.e., the die is fair.

Your friend has picked at random a multi-sided die. (You do not know what sided-die he has picked.) He then rolled the die and told you the result. You are to estimate the result when he rolls that same die again.

If the result is a 3 then the estimate of the next roll of the same die is 2.853:

A	B	C	D	E	F
Type of Die	A Priori Chance of This Type of Die	Chance of the Observation	Prob. Weight = Product of Columns B & C	Posterior Chance of This Type of Die	Mean Die Roll
4-sided	0.600	0.250	0.1500	70.6%	2.5
6-sided	0.300	0.167	0.0500	23.5%	3.5
8-sided	0.100	0.125	0.0125	5.9%	4.5
Overall			0.2125	1.000	**2.853**

Example 4.2.3: If instead a 6 is rolled, what is the estimate of the next roll of the same die?

[*Solution:* The estimate of the next roll of the same die is 3.700:

A	B	C	D	E	F
Type of Die	A Priori Chance of This Type of Die	Chance of the Observation	Prob. Weight = Product of Columns B & C	Posterior Chance of This Type of Die	Mean Die Roll
4-sided	0.600	0.000	0.0000	0.0%	2.5
6-sided	0.300	0.167	0.0500	80.0%	3.5
8-sided	0.100	0.125	0.0125	20.0%	4.5
Overall			0.0625	1.000	**3.700**

For this example we get the following set of estimates corresponding to each possible observation:

Observation	1	2	3	4	5	6	7	8
Bayesian Estimate	2.853	2.853	2.853	2.853	3.7	3.7	4.5	4.5

Note that while in this simple example the posterior estimates are the same for a number of different observations, this is not usually the case.]

Relation of Bayesian Analysis and Bühlmann Credibility

As discussed in section 3.2 on Bühlmann Credibility, in the multi-sided dice example $K = \text{EPV}/\text{VHM} = 2.15/.45 = 4.778 = 43/9$. For one observation, $Z = 1/(1+4.778) = .1731 = 9/52 = .45/(.45+2.15)$.

The Bühlmann Credibility estimate is a linear function of the observation:

Observation	1	2	3	4	5	6	7	8
New Estimate	2.6538	2.8269	3	3.1731	3.3462	3.5193	3.6924	3.8655

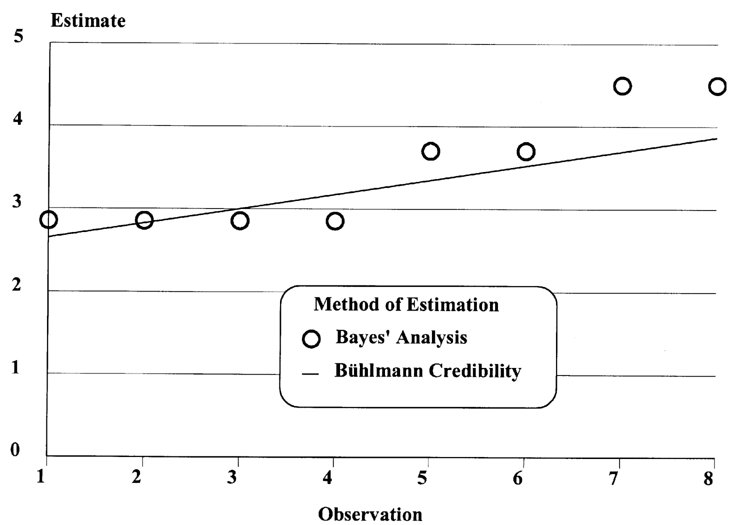
The above results of applying Bühlmann Credibility differ from those obtained for Bayesian Analysis.

Observation	1	2	3	4	5	6	7	8
Bühlmann Credibility Estimate	2.6538	2.8269	3	3.1731	3.3462	3.5193	3.6924	3.86552
Bayesian Analysis Estimate	2.853	2.853	2.853	2.853	3.7	3.7	4.5	4.5

We note that in this case the line formed by the Bühlmann Credibility estimates seems to approximate the Bayesian Analysis Estimates as shown in Graph 3. In general it turns out that the Bühlmann Credibility Estimates are the weighted least squares line fit to the Bayesian Estimates.

GRAPH 3

MULTI-SIDED DIE EXAMPLE



4.2. Exercises

Use the following information for the next seven questions:

There are three types of risks. Assume 60% of the risks are of Type A, 25% of the risks are of Type B, and 15% of the risks are of Type C. Each risk has either one or zero claims per year. A risk is selected at random.

Type of Risk	A Priori Chance of Type of Risk	Chance of a Claim
A	60%	20%
B	25%	30%
C	15%	40%

4.2.1. What is the overall mean annual claim frequency?

4.2.2. You observe no claim in a year. What is the probability that the risk you are observing is of Type A?

4.2.3. You observe no claim in a year. What is the probability that the risk you are observing is of Type B?

4.2.4. You observe no claim in a year. What is the probability that the risk you are observing is of Type C?

4.2.5. You observe no claim in a year. What is the expected annual claim frequency from the same risk?

4.2.6. You observe one claim in a year. What is the expected annual claim frequency from the same risk?

4.2.7. You observe a single risk over five years. You observe 2 claims in 5 years. What is the expected annual claim frequency from the same risk?

4.2.8. Let X_1 be the outcome of a single trial and let $E[X_2 \mid X_1]$ be the expected value of the outcome of a second trial. You are given the following information:

Outcome = T	$P(X_1 = T)$	Bayesian Estimate for $E[X_2 \mid X_1 = T]$
1	5/8	1.4
4	2/8	3.6
16	1/8	—

Determine the Bayesian estimate for $E[X_2 \mid X_1 = 16]$.

Use the following information for the next two questions:

There are two types of urns, each with many balls labeled \$1,000 and \$2,000.

Type of Urn	A Priori Chance of This Type of Urn	Percentage of \$1,000 Balls	Percentage of \$2,000 Balls
I	80%	90%	10%
II	20%	70%	30%

4.2.9. You pick an Urn at random and pick one ball. If the ball is \$2,000, what is the expected value of the next ball picked from that same urn?

4.2.10. You pick an Urn at random (80% chance it is of Type I) and pick three balls, returning each ball to the Urn before the next pick. If two of the balls were \$1,000 and one of the balls was \$2,000, what is the expected value of the next ball picked from that same urn?

4.2.11. You are given the following information:

- There are three types of risks
- The types are homogeneous, every risk of a given type has the same Poisson frequency process:

Type	Portion of Risks in This Type	Average (Annual) Claim Frequency
1	70%	.4
2	20%	.6
3	10%	.8

A risk is picked at random and we do not know what type it is. For this randomly selected risk, during one year there are 3 claims. Use Bayesian Analysis to predict the future claim frequency of this same risk.

Use the following information for the next two questions:

There are three marksmen, each of whose shots are Normally Distributed (in one dimension) with means and standard deviations:

Risk	Mean	Standard Deviation
A	10	3
B	20	5
C	30	15

4.2.12. A marksman is chosen at random. If you observe two shots at 10 and 14, what is the chance that it was marksman B?

4.2.13. A marksman is chosen at random. If you observe two shots at 10 and 14, what is the Bayesian Estimate of the next shot from the same marksman?

4.3. Continuous Prior and Posterior Distributions

So far in this section on Bayesian analysis our examples included discrete probability distributions. Now we'll look at continuous distributions.

Let random variables X and Y have the joint p.d.f. (probability density function) $f(x,y)$. The marginal p.d.f. for X is $f_X(x) = \int_{-\infty}^{\infty} f(x,y)dy$ and the marginal p.d.f. for Y is $f_Y(y) = \int_{-\infty}^{\infty} f(x,y)dx$.[48] The conditional p.d.f. for Y given X is:

$$f_Y(y \mid x) = \frac{f(x,y)}{f_X(x)} \tag{4.3.1}$$

Note the correspondence between this formula and formula (4.1.1): $P[A \mid B] = P[A \text{ and } B]/P[B]$. Formula (4.3.1) applies to both continuous and discrete p.d.f.s, or a combination of the two.

In insurance applications it is common for one of the random variables, call it X, to be a discrete random variable and a second variable, θ, to be continuous. X may be the number of claims.[49] The other random variable, θ, is a parameter of the distribution of X. The conditional p.d.f. of X given θ is denoted $f_X(x \mid \theta)$ and the p.d.f. for θ is $f_\theta(\theta)$. The joint p.d.f. for X and θ is:

$$f(x,\theta) = f_X(x \mid \theta) f_\theta(\theta) \tag{4.3.2}$$

The p.d.f. $f_\theta(\theta)$ is the **prior distribution** of θ. It may represent our initial guess about the distribution of some characteristic within a population, for example expected claim frequency.

If we select a risk at random from the population and observe the number of claims, x, then we can update our estimate of θ for this risk. Using formula (4.3.1) and replacing the joint p.d.f.

[48] If one of the random variables, say Y, is discrete, then the integral is replaced with a summation and the marginal p.d.f. for X becomes: $f_X(x) = \sum_i f(x,y_i)$.

[49] If X represents claim sizes, then X is usually continuously distributed. Examples are given in the Appendix.

of X and θ with (4.3.2) produces:

$$f_\theta(\theta \mid X = x) = f_X(x \mid \theta) f_\theta(\theta)/f_X(x) \tag{4.3.3}$$

$f_\theta(\theta \mid X = x)$ is the **posterior distribution** of θ for the selected risk. We have used information about the risk to adjust our estimate of the distribution of θ. Equation 4.4.3 is just another form of Bayes Theorem, equation 4.2.1.

Example 4.3.1: The probability of exactly one claim during a year for any insured in a population is θ. The probability of no claims is $(1-\theta)$. The probability of a claim, θ, varies within the population with a p.d.f. given by a uniform distribution:

$$f_\theta(\theta) = 1 \qquad \text{if} \quad 0 \leq \theta \leq 1, \quad 0 \text{ otherwise.}$$

An insured, Bob, is selected at random from the population. Bob is observed to have a claim during the next year, i.e. $X = 1$. Calculate the posterior density $f_\theta(\theta \mid X = 1)$ for Bob.

[*Solution:* Calculate the marginal p.d.f. for X evaluated at $X = 1$: $f_X(1) = \int_0^1 f(1,\theta)d\theta = \int_0^1 f_X(1 \mid \theta) f_\theta(\theta) d\theta$. Since $f_X(1 \mid \theta) = \theta$ and $f_\theta(\theta) = 1$ over [0,1], then $f_X(1) = \int_0^1 \theta d\theta = 1/2$. So $f_\theta(\theta \mid X = 1) = f_X(X = 1 \mid \theta) f_\theta(\theta)/f_X(1) = (\theta)(1)/(1/2) = 2\theta$.]

Note that in this example the distribution has shifted from being uniform over the interval [0,1] to one that rises linearly from 0 at the left endpoint to 2 at the right endpoint. The fact that Bob had a claim has shifted the weight of the p.d.f. of θ to the right.

Knowing the conditional p.d.f. allows the calculation of the conditional expectation:

$$E[X \mid Y] = \int_{-\infty}^{\infty} x f_X(x \mid y) dx \tag{4.3.4}$$

And, the expectation of X is:

$$E[X] = E_Y[E_X[X \mid Y]] = \int_{-\infty}^{\infty} \left\{ \int_{-\infty}^{\infty} x f_X(x \mid y)dx \right\} f_Y(y)dy \tag{4.3.5}$$

Example 4.3.2: Assuming the information from Example 4.3.1, calculate the following:

1. The expectation for the number of claims in a year for an insured selected at random from the population.
2. The expectation for the number of claims in a year for an insured who was randomly selected from the population and then had one claim during an observation year.

[*Solution:* (1) The distribution of X, the number of claims, is discrete given θ. $E_X[X \mid \theta] = (1)(\theta) + (0)(1-\theta) = \theta$. Since $f_\theta(\theta) = 1$ for $0 < \theta < 1$, $E[X] = E_\theta[E_X[X \mid \theta]] = E_\theta[\theta] = \int_0^1 \theta f_\theta(\theta)d\theta = \int_0^1 \theta d\theta = 1/2$.

(2) As in (1), $E_X[X \mid \theta] = \theta$. But from example 4.3.1, the posterior p.d.f. of θ given that $X = 1$ is $f_\theta(\theta \mid X = 1) = 2\theta$. So, $E[X] = E_\theta[E_X[X \mid \theta]] = E_\theta[\theta] = \int_0^1 \theta(2\theta)d\theta = 2/3$.]

Bayesian Interval Estimates

By the use of Bayes Theorem one obtains an entire posterior distribution. Rather than just using the mean of that posterior distribution in order to get a point estimate, one can use the posterior density function to estimate the posterior chance that the quantity of interest is in a given interval. This is illustrated in the next example.

Example 4.3.3: Assume the information from example 4.3.1. Given that Bob had one claim in one year, what is the posterior estimate that he has a θ parameter less than .2?

[*Solution:* From example 4.3.1, the posterior density function for θ is: $f_\theta(\theta \mid X = 1) = 2\theta$.

The posterior chance that θ is in the interval $[0,.2]$ is the integral from 0 to .2 of the posterior density:

$$\int_0^{.2} 2\theta\, d\theta = \theta^2 \Big]_{p=0}^{p=.2} = .2^2 = .04.]$$

The idea of a continuous distribution of types of risks is developed for the particular case of the Gamma-Poisson conjugate prior in Section 5.

4.3. Exercises

For the next two problems, assume the following information:

The probability of exactly one claim during a year for any insured in a population is θ. The probability of no claims is $(1-\theta)$. The probability of a claim, θ, varies within the population with a p.d.f. given by a uniform distribution:

$$f_\theta(\theta) = 1 \qquad \text{if} \quad 0 \le \theta \le 1, \quad 0 \text{ otherwise.}$$

An insured, Bob, is selected at random from the population.

4.3.1. Bob is observed for two years after being randomly selected from the population and has a claim in each year. Calculate the posterior density function of θ for Bob.

4.3.2. Bob is observed for three years after being randomly selected from the population and has a claim in each year. Calculate the posterior density function of θ for Bob.

Use the following information for the next two questions:

- The probability of y successes in n trials is given by a Binomial distribution with parameters n and p.
- The prior distribution of p is uniform on $[0,1]$.
- One success was observed in three trials.

4.3.3. What is the Bayesian estimate for the probability that the unknown parameter p is in the interval $[.3, .4]$?

4.3.4. What is the probability that a success will occur on the fourth trial?

4.3.5. A number x is randomly selected from a uniform distribution on the interval $[0, 1]$. Three independent Bernoulli trials are performed with probability of success x on each trial. All three are successes. What is the posterior probability that x is less than 0.9?

4.3.6. You are given the following:

- The probability that a single insured will produce 0 claims during the next exposure period is $e^{-\theta}$.
- θ varies by insured and follows a distribution with density function

$$f(\theta) = 36\theta e^{-6\theta}, \qquad 0 < \theta < \infty.$$

Determine the probability that a randomly selected insured will produce 0 claims during the next exposure period.

4.3.7. Let N be the random variable that represents the number of claims observed in a one-year period. N is Poisson distributed with a probability density function with parameter θ:

$$P[N = n \mid \theta] = e^{-\theta}\theta^n/n!, \qquad n = 0, 1, 2, \ldots$$

The probability of observing no claims in a year is less than .450. Which of the following describe possible probability distributions for θ?

1. θ is uniformly distributed on $(0, 2)$.
2. The probability density function of θ is $f(\theta) = e^{-\theta}$ for $\theta > 0$.
3. $P[\theta = 1] = 1$ and $P[\theta \neq 1] = 0$.

5. CONJUGATE PRIORS

Conjugate prior distributions have elegant mathematical properties that make them valuable in actuarial applications. The Gamma-Poisson is the most important of these for casualty actuarial work. A study of the Gamma-Poisson model is valuable to the understanding of Bayesian analysis and Bühlmann Credibility.[50]

5.1. Gamma Function and Distribution

The quantity $x^{\alpha-1}e^{-x}$ is finite for $x \geq 0$ and $\alpha \geq 1$. Since it declines quickly to zero as x approaches infinity, its integral from zero to infinity exists. This is the much studied and tabulated (complete) Gamma Function.[51]

$$\Gamma(\alpha) = \int_{t=0}^{\infty} t^{\alpha-1}e^{-t}dt = \lambda^{\alpha}\int_{t=0}^{\infty} t^{\alpha-1}e^{-\lambda t}dt \qquad \text{for} \quad \alpha \geq 0, \quad \lambda \geq 0. \tag{5.1.1}$$

It can be proven that:[52] $\Gamma(\alpha) = (\alpha-1)\Gamma(\alpha-1)$.

For integral values of α, $\Gamma(\alpha) = (\alpha-1)!$, and $\Gamma(1) = 1$, $\Gamma(2) = 1! = 1$, $\Gamma(3) = 2! = 2$, $\Gamma(4) = 3! = 6$, $\Gamma(5) = 4! = 24$, etc.

Integrals involving e^{-x} and powers of x can be written in terms of the Gamma function:

$$\int_{t=0}^{\infty} t^{\alpha-1}e^{-\lambda t}dt = \Gamma(\alpha)\lambda^{-\alpha}. \tag{5.1.2}$$

Equation 5.1.2 is very useful for working with anything involving the Gamma distribution, for example the Gamma-Poisson process as will be seen below. (It follows from the definition of the Gamma function and a change of variables.) The probability

[50]This is a special case of a general result for conjugate priors of members of "linear exponential families." This general result is beyond the scope of this chapter.
[51]See for example, Handbook of Mathematical Functions, Milton Abramowitz, et. al., National Bureau of Standards, 1964.
[52]Use integration by parts: $\int u\,dv = uv - \int v\,du$ with $u = t^{\alpha-1}$ and $dv = e^{-t}dt$.

density of the Gamma Distribution is: $f(x) = \lambda^\alpha x^{\alpha-1}e^{-\lambda x}/\Gamma(\alpha)$. Since the probability density function must integrate to unity, we must have the above relationship. This is a useful way to remember Equation 5.1.2.

Example 5.1.1: In terms of its parameters, α and λ, what is the mean of a Gamma distribution?

[*Solution:* The mean is the expected value of x.

$$\int_0^\infty xf(x)dx = \int_0^\infty xx^{\alpha-1}e^{-\lambda x}dx(\lambda^\alpha/\Gamma(\alpha))$$

$$= \int_0^\infty x^\alpha e^{-\lambda x}dx(\lambda^\alpha/\Gamma(\alpha)) = \{\Gamma(\alpha+1)/\lambda^{a+1}\}\{\lambda^\alpha/\Gamma(\alpha)\}$$

$$= \{\Gamma(\alpha+1)/\Gamma(\alpha)\}/\lambda = \alpha/\lambda.]$$

Example 5.1.2: In terms of its parameters, α and λ, what is the *n*th moment of a Gamma distribution?

[*Solution:* The n^{th} moment is the expected value of x^n.

$$\int_0^\infty x^n f(x)dx = \int_0^\infty x^n x^{\alpha-1}e^{-\lambda x}dx(\lambda^\alpha/\Gamma(\alpha))$$

$$= \int_0^\infty x^{\alpha+n-1}e^{-\lambda x}dx(\lambda^\alpha/\Gamma(\alpha)) = \{\Gamma(\alpha+n)/\lambda^{\alpha+n}\}\{\lambda^\alpha/\Gamma(\alpha)\}$$

$$= \{\Gamma(\alpha+n)/\Gamma(\alpha)\}/\lambda^n = (\alpha+n-1)(\alpha+n-2)\ldots(\alpha)/\lambda^n.]$$

5.1. Exercises

5.1.1. What is the value of the integral from zero to infinity of x^5e^{-8x}?

5.1.2. What is the density at $x = 8$ of a Gamma distribution with $\alpha = 3$ and $\lambda = .10$?

5.1.3. Using the results of example 5.1.2, show that the variance of the Gamma distribution is α/λ^2.

5.1.4. If $\alpha = 3.0$ and $\lambda = 1.5$, what are the mean and variance of the Gamma distribution?

5.2. The Gamma-Poisson Model

The Gamma-Poisson model has two components: (1) a Poisson distribution which models the number of claims for an insured with a given claims frequency, and (2) a Gamma distribution to model the distribution of claim frequencies within a population of insureds. As in previous sections the goal is to use observations about an insured to infer future expectations.

Poisson Distribution

We'll assume that the number of claims for an insured is Poisson distributed and that the average number of claims in a year is μ. The probability of having n claims in a year is given by:

$$P[n \mid \mu] = \mu^n e^{-\mu}/n!. \tag{5.2.1}$$

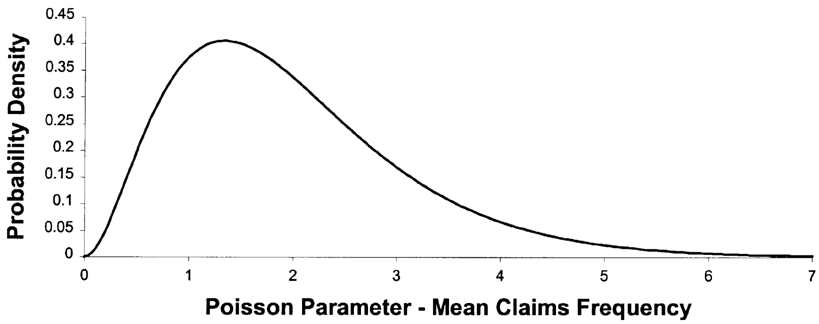
μ is the mean annual number of claims, i.e. $E[n] = \mu$. Any particular insured within the population is assumed to have a μ that remains constant over the time period of interest. However, the estimation of μ is the challenge since μ's may vary from risk to risk. You do not know μ for a risk selected at random.

Example 5.2.1: The number of claims has a Poisson distribution with parameter $\mu = 2$. Calculate the separate probabilities of exactly 0, 1, and 2 claims. Calculate the mean and variance of the distribution.

[*Solution:* $f(n) = \mu^n e^{-\mu}/n!$. Since $\mu = 2.0$, $f(0) = 2^0 e^{-2}/0! = .135$, $f(1) = 2^1 e^{-2}/1! = .271$, and $f(2) = 2^2 e^{-2}/2! = .271$. For the Poisson distribution, mean = variance = $\mu = 2.0$.]

GRAPH 4

GAMMA DISTRIBUTION



Prior Gamma Distribution

The mean claim frequencies for insureds within the population are assumed to be Gamma distributed with probability density function:

$$f(\mu) = \lambda^{\alpha}\mu^{\alpha-1}e^{-\lambda\mu}/\Gamma(\alpha), \qquad \text{for} \quad \mu > 0 \quad \text{and} \quad \alpha, \lambda > 0. \tag{5.2.2}$$

The random variable is μ and the parameters that determine the shape and scale of the distribution are α and λ. The mean of the distribution is α/λ. So, the average claims frequency across all insureds in the population is α/λ. The variance is α/λ^2. $f(\mu)$ defines the **prior distribution** of the mean claim frequencies.

Let's consider a particular case of the Gamma p.d.f. with parameters $\alpha = 3$, and $\lambda = 1.5$. The p.d.f. for μ is:

$$f(\mu) = 1.5^3\mu^{3-1}e^{-1.5\mu}/\Gamma(3) = 1.6875\mu^2 e^{-1.5\mu}, \qquad \text{for} \quad \mu > 0.$$

This p.d.f. is shown in Graph 4.

Gamma-Poisson Mixed Distribution

Suppose that an insured is selected at random from the population of insureds. What is the distribution of the number of claims for the insured? Or, equivalently, what is the probability that you will see exactly n claims in the next time period?

The probability of exactly n claims for an insured with mean claim frequency μ is $P[n \mid \mu] = \mu^n e^{-\mu}/n!$. But, μ varies across the population. We need to calculate a weighted average of the $P[n \mid \mu]$'s. What weight should we use? Use the p.d.f. (5.2.2), the relative weights of the μ's within the population. After applying weight $f(\mu)$, then we sum over all insureds by taking the integral.

The number of claims for an insured selected at random from the population has a **Gamma-Poisson mixed distribution** defined by the p.d.f.:

$$
\begin{aligned}
g(n) &= \int_0^\infty P[n \mid \mu] f(\mu) d\mu \\
&= \int_0^\infty (\mu^n e^{-\mu}/n!) f(\mu) d\mu = \int_0^\infty (\mu^n e^{-\mu}/n!)(\lambda^\alpha \mu^{\alpha-1} e^{-\lambda\mu}/\Gamma(\alpha)) d\mu \qquad (5.2.3\text{a}) \\
&= (\lambda^\alpha / n!\Gamma(\alpha)) \int_0^\infty \mu^{n+\alpha-1} e^{-(\lambda+1)\mu} d\mu \\
&= \{\lambda^\alpha / n!\Gamma(\alpha)\}\{\Gamma(n+\alpha)(\lambda+1)^{-(n+\alpha)}\}^{53} \\
&= \{\Gamma(n+\alpha)/n!\Gamma(\alpha)\}\{\lambda/(\lambda+1)\}^\alpha \{1/(\lambda+1)\}^n \qquad (5.2.3\text{b}) \\
&= \{(n+\alpha-1)!/n!(\alpha-1)!\}\{\lambda/(\lambda+1)\}^\alpha \{1-\lambda/(\lambda+1)\}^n \\
&= \binom{n+\alpha-1}{n} p^\alpha (1-p)^n \qquad (5.2.3\text{c})
\end{aligned}
$$

Note that we have substituted p for $\lambda/(\lambda+1)$ in (5.2.3c).

[53] In this derivation we used equation (5.1.2): $\Gamma(\alpha)\lambda^{-\alpha} = \int_0^\infty t^{\alpha-1} e^{-\lambda t} dt$. Through substitution of t for $(\lambda+1)\mu$, it can be shown that $\int_0^\infty \mu^{n+\alpha-1} e^{-(\lambda+1)\mu} d\mu = \Gamma(n+\alpha)(\lambda+1)^{-(n+\alpha)}$.

Thus the prior mixed distribution is in the form of the Negative Binomial distribution with $k = \alpha$ and $p = \lambda/(\lambda+1)$:

$$\binom{n+k-1}{n} p^k (1-p)^n \tag{5.2.3d}$$

The Negative Binomial distribution evaluated at n is the probability of seeing exactly n claims in the next year for an insured selected at random from the population. The mean of the negative binomial is $k(1-p)/p$ and the variance is $k(1-p)/p^2$. (See the Appendix.) In terms of α and λ, the mean is α/λ and the variance is $\alpha(\lambda+1)/\lambda^2$.

Example 5.2.2: The number of claims has a Negative Binomial distribution with parameters $k = \alpha$ and $p = \lambda/(\lambda+1)$. Assume $\alpha = 3.0$ and $\lambda = 1.5$. Calculate the separate probabilities of exactly 0, 1, and 2 claims. Also calculate the mean and variance of the distribution.

[*Solution:* First, $p = \lambda/(\lambda+1) = 1.5/(1.5+1) = .60$ and $k = \alpha = 3.0$. Next $f(n) = \{(n+\alpha-1)!/(n!(\alpha-1)!)\}p^\alpha(1-p)^n = \{(n+3-1)!/(n!(3-1)!)\}(.6^3)(1-.6)^n$. So, $f(0) = \{(0+2)!/(0!(2)!)\}.6^3(.4)^0 = .216$, $f(1) = \{(1+2)!/(1!(2)!)\}.6^3(.4)^1 = .259$, and $f(2) = \{(2+2)!/(2!(2)!)\}.6^3(.4)^2 = .207$. The mean $= k(1-p)/p = \alpha/\lambda = 2.0$ and variance $= k(1-p)/p^2 = \alpha(\lambda+1)/\lambda^2 = 3(1.5+1)/1.5^2 = 3.33$.]

Compare example 5.2.2 with example 5.2.1. Even though the two distributions have the same mean 2.0, the variance of the Negative Binomial is larger.[54] The uncertainty introduced by the random selection of an insured from the population has increased the variance.

5.2. *Exercises*

5.2.1. For an insurance portfolio the distribution of the number of claims a particular policyholder makes in a year

[54]For Negative Binomial, variance $>$ mean. For Poisson, variance = mean. For Binomial, variance $<$ mean.

is Poisson with mean μ. The μ-values of the policyholders follow the Gamma distribution, with parameters $\alpha = 4$, and $\lambda = 9$. Determine the probability that a policyholder chosen at random will experience 5 claims over the next year.

5.2.2. The number of claims X for a given insured follows a Poisson distribution, $P[X = x] = \theta^x e^{-\theta}/x!$. Over the population of insureds the expected annual mean of the Poisson distribution follows the distribution $f(\theta) = 9\theta e^{-3\theta}$ over the interval $(0,\infty)$. An insured is selected from the population at random. What are the mean and variance for the number of claims for the selected insured?

5.2.3. Assume that random variable X is distributed according to a Negative Binomial with parameters $k = 2$ and $p = 0.6$. What is the probability that the observed value of X is greater than 2?

5.2.4. Assume the following information:

1. the claim count N for an individual insured has a Poisson distribution with mean λ; and
2. λ is uniformly distributed between 1 and 3.

Find the probability that a randomly selected insured will have no claims.

5.2.5. Prove each of the following:

1. For the Binomial distribution, the mean is greater than or equal to the variance.
2. For the Negative Binomial, the mean is less than or equal to the variance.

If the means are equal for a Binomial distribution, Poisson distribution, and a Negative Binomial distribution, rank the variances by size for the three distributions.

5.3. Bayesian Analysis on the Gamma-Poisson Model

The Gamma p.d.f. $f(\mu)$ defined in section 5.2 is the prior distribution for the mean annual claims frequency μ for a risk selected at random from the population. But, the p.d.f. can be updated using Bayesian analysis after observing the insured. The distribution of μ subsequent to observations is referred to as the **posterior distribution**, as opposed to the prior distribution.

Suppose that the insured generates C claims during a one-year observation period. We want to calculate the posterior distribution for μ given this information: $f(\mu \mid n = C)$. Bayes Theorem stated in terms of probability density functions is:[55]

$$f(\mu \mid n = C) = P[C \mid \mu]f(\mu)/P(C) \tag{5.3.1}$$

We have all of the pieces on the right hand side. They are formulas (5.2.1), (5.2.2), and (5.2.3b). Putting them all together:

$$f(\mu \mid n = C) = [\mu^C e^{-\mu}/C!][\lambda^\alpha \mu^{\alpha-1} e^{-\lambda\mu}/\Gamma(\alpha)]/$$
$$[\{\Gamma(C+\alpha)/C!\Gamma(\alpha)\}\{\lambda/(\lambda+1)\}^\alpha\{1/(\lambda+1)\}^C]$$

Through cancellations and combining terms, this can be simplified to:

$$f(\mu \mid n = C) = (\lambda+1)^{\alpha+C}\mu^{\alpha+C-1}e^{-(\lambda+1)\mu}/\Gamma(\alpha+C) \tag{5.3.2}$$

Substituting $\alpha' = \alpha + C$ **and** $\lambda' = \lambda + 1$**, yields**:

$$f(\mu \mid n = C) = \lambda'^{\alpha'}\mu^{\alpha'-1}e^{-\lambda'\mu}/\Gamma(\alpha') \tag{5.3.3}$$

This is the posterior distribution for μ and it is also a Gamma distribution.

The fact that the posterior distribution is of the same form as the prior distribution is why the Gamma is called a Conjugate Prior Distribution for the Poisson.

---

[55] See Section 3.3.

Example 5.3.1: The distribution of annual claims frequencies μ within a population of insureds is Gamma with parameters α = 3.0 and λ = 1.5. An insured is selected at random.

1. What is the expected value of μ?
2. If the insured is observed to have a total of 0 claims during a one-year observation period, what is the expected value of μ for the insured?
3. The insured is observed to have a total of 5 claims during a one-year observation period, what is the expected value of μ for the insured?

[*Solution:* (1) $E[\mu] = \alpha/\lambda = 3.0/1.5 = 2.0$. (2) $E[\mu \mid 0$ claims in one year$] = (\alpha + 0)/(\lambda + 1) = (3 + 0)/(1.5 + 1) = 1.2$. (3) $E[\mu \mid 5$ claims in one year$] = (\alpha + 5)/(\lambda + 1) = (3 + 5)/(1.5 + 1) = 3.2$.]

In example 5.3.1, prior to any observations of the insured, our estimate of the expected claim frequency μ is just the population average 2.0 claims per year. If we observe 0 claims in one-year, then we lower our estimate to 1.2 claims per year. On the other hand, if we observe 5 claims in one year, we raise our estimate to 3.2.

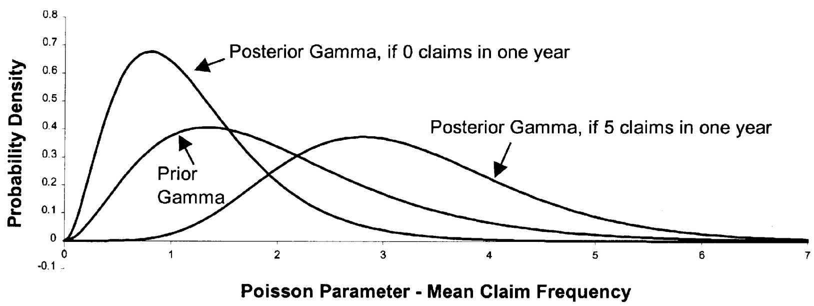
Although the distribution of μ remains Gamma as more information is gathered about the risk, the shape of the distribution changes. Graph 5 shows as in example 5.3.1, (1) a prior distribution with α = 3.0 and λ = 1.5, (2) the posterior distribution after observing 0 claims in one-year, and (3) the posterior distribution after observing 5 claims in one-year.

Multiple Years of Observation

Suppose we now observe the insured for two years and see C_1 claims in the first year and C_2 claims in the second year. What is the new posterior distribution? The posterior distribution $f(\mu \mid n = C_1)$ after the first observation year becomes the new prior distribution at the start of the second observation year. After observing C_2 claims during the second year, the posterior

GRAPH 5

GAMMA DISTRIBUTION



distribution is again a Gamma distribution, but with parameters $\alpha'' = \alpha' + C_2 = \alpha + C_1 + C_2$ and $\lambda'' = \lambda' + 1 = \lambda + 1 + 1$. Continuing the observations for a total of Y years and total observed claims of $C = C_1 + C_2 + \cdots + C_Y$ produces a final posterior distribution that is still Gamma but with parameters:

$$\hat{\alpha} = \alpha + C$$

$$\hat{\lambda} = \lambda + Y$$

The mean of the posterior distribution is $(\alpha + C)/(\lambda + Y)$. This is $E[\mu \mid C$ claims in Y years$]$ and the expected value of the annual claims frequency for the insured.

Thus for the **Gamma-Poisson the posterior density function is also a Gamma. This posterior Gamma has a first parameter equal to the prior first parameter plus the number of claims observed. The second parameter equals the prior second parameter plus the number of exposures (usually years) observed.**

Example 5.3.2: The distribution of annual claims frequencies μ within a population of insureds is Gamma with parameters $\alpha = 3.0$ and $\lambda = 1.5$. An insured is selected at random and then observed for two years. The insured has two claims during the first year and four claims during the second year. What are the parameters of the posterior Gamma distribution for this risk?

[*Solution:* $C = 2 + 4 = 6$ and $Y = 2$, so $\hat{\alpha} = 3 + 6 = 9$ and $\hat{\lambda} = 1.5 + 2 = 3.5$.]

Predictive Distribution

The distribution of the number of claims for the selected insured still follows a Negative Binomial. That's because the posterior distribution is still Gamma. But, in formula (5.2.3c) α is replaced by $\alpha + C$ and p is replaced $p = (\lambda + Y)/(\lambda + 1 + Y)$. This posterior mixed distribution is referred at the **predictive distribution.**

Example 5.3.3: The distribution of annual claims frequencies μ within a population of insureds is Gamma with parameters $\alpha = 3.0$ and $\lambda = 1.5$. An insured is selected at random.

1. Prior to any observations of the insured, what is the probability that the insured will have two or more claims in a year?

Suppose that the insured is observed to have two claims during the first year and four claims during the second year.

2. What is the probability that the insured will have two or more claims in a year?

[*Solution:* (1) With $\alpha = 3.0$ and $\lambda = 1.5$, the parameters for the Negative Binomial distribution are $k = 3$ and $p = 1.5/(1.5 + 1) = .6$. (See formula (5.2.3d).) Then,

$$P[0 \text{ claims}] = \binom{2}{0} .6^3 .4^0 = .216 \qquad \text{and}$$

$$P[1 \text{ claim}] = \binom{3}{1} .6^3 .4^1 = .2592.$$

So, $P[2 \text{ or more claims}] = 1 - .216 - .2592 = .5248$.

(2) The new parameters for the Gamma posterior distribution are $\hat{\alpha} = 3 + 6 = 9$ and $\hat{\lambda} = 1.5 + 2 = 3.5$. The parameters for the predictive Negative Binomial distribution are $k = \hat{\alpha} = 9$ and $p = 3.5/(1 + 3.5) = .778$. Using formula (5.2.3d),

$$P[0 \text{ claims}] = \binom{8}{0}.778^9.222^0 = .104 \qquad \text{and}$$

$$P[1 \text{ claim}] = \binom{9}{1}.778^9.222^1 = .209.$$

So, $P[2 \text{ or more claims}] = 1 - .104 - .209 = .687$. The fact that we observed 6 claims in two years has raised our estimate of the probability of having two or more claims in a year versus our estimate for someone randomly selected from the population without any additional information.]

The Gamma-Poisson is one example of conjugate prior distributions. There are many other conjugate priors. Examples include the Beta-Bernoulli for frequency and the Normal-Normal for severity.

Figure 4 shows the relationships between the Gamma Prior, Gamma Posterior, Negative Binomial mixed distribution, and Negative Binomial predictive distribution for the Gamma-Poisson frequency process.

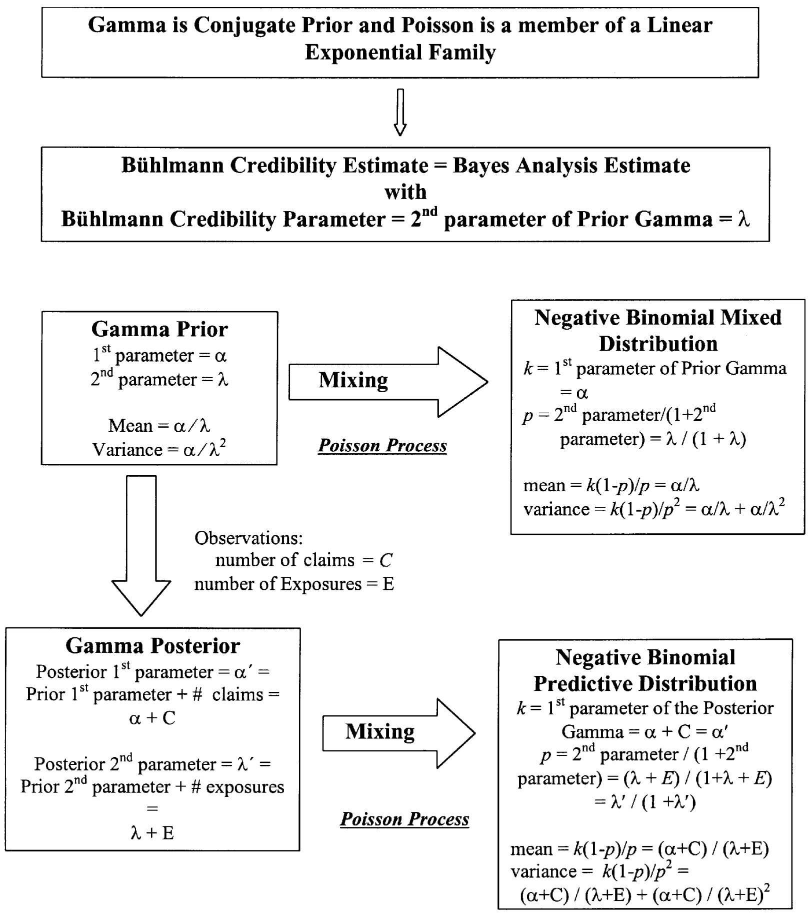
5.3. Exercises

5.3.1. The number of claims is distributed according to a Gamma-Poisson mixed distribution. The prior Gamma has parameters $\alpha = 4$ and $\lambda = 2$. Over a three-year period, 5 claims were observed. Calculate the parameters α' and λ' of the posterior Gamma distribution.

5.3.2. Let the likelihood of a claim be given by a Poisson distribution with parameter θ. The prior density function of θ is given by $f(\theta) = 32\theta^2 e^{-4\theta}$. You observe 1 claim in 2 years. What is the posterior density function of θ?

FIGURE 4

GAMMA-POISSON FREQUENCY PROCESS



Poisson parameters of individuals making up the entire portfolio are distributed via a Gamma Distribution with parameters α and $\lambda : f(x) = \lambda^{\alpha} x^{\alpha-1} e^{-\lambda x} / \Gamma(\alpha)$.

5.3.3. The number of claims X for a given insured follows a Poisson distribution, $P[X = x] = \theta^x e^{-\theta}/x!$. The expected annual mean of the Poisson distribution over the population of insureds follows the distribution $f(\theta) = e^{-\theta}$ over the interval $(0, \infty)$. An insured is selected from the population at random. Over the last year this particular insured had no claims. What is the posterior density function of θ for the selected insured?

5.3.4. An automobile insurer entering a new territory assumes that each individual car's claim count has a Poisson distribution with parameter μ and that μ is the same for all cars in the territory for the homogeneous class of business that it writes. The insurer also assumes that μ has a gamma distribution with probability density function

$$f(\mu) = \lambda^\alpha \mu^{\alpha-1} e^{-\lambda\mu}/\Gamma(\alpha)$$

Initially, the parameters of the gamma distribution are assumed to be $\alpha = 50$ and $\lambda = 500$. During the subsequent two-year period the insurer covered 750 and 1100 cars for the first and second years, respectively. The insurer incurred 65 and 112 claims in the first and second years, respectively. What is the coefficient of variation of the posterior gamma distribution? (The coefficient of variation is the standard deviation divided by the mean.)

5.3.5. The likelihood of a claim is given by a Poisson distribution with parameter θ. The prior density function of θ is given by $f(\theta) = 32\theta^2 e^{-4\theta}$. A risk is selected from a population and you observe 1 claim in 2 years. What is the probability that the mean claim frequency for this risk falls in the interval $[1, 2]$? (Your answer can be left in integral form.)

5.3.6. You are given the following:

- You are trying to estimate the average number of claims per exposure, μ, for a group of insureds.
- Number of claims follows a Poisson distribution.

- Prior to the first year of coverage, μ is assumed to have the Gamma distribution $f(\mu) = 1000^{150}\mu^{149}e^{-1000\mu}/\Gamma(150)$, $\mu > 0$.
- In the first year, 300 claims are observed on 1,500 exposures.
- In the second year, 525 claims are observed on 2,500 exposures.

After two years, what is the Bayesian probability estimate of $E[\mu]$?

5.4. Bühlmann Credibility in the Gamma-Poisson Model

As in section 5.3, an insured is selected at random from a population with a Gamma distribution of average annual frequencies. Using Bühlmann Credibility we want to estimate the expected annual claims frequency. To do this, we need to calculate the Expected Value of the Process Variance (EPV) and the Variance of the Hypothetical Means (VHM).

Expected Value of the Process Variance

The means μ in the population are distributed according to the Gamma distribution shown in (5.2.2) with parameters α and λ. The mean of the Gamma is α/λ, so $E[\mu] = \alpha/\lambda$. For the Poisson distribution, the means and process variances are equal. So, EPV = E[Process Variance] = E[Mean] = $E[\mu] = \alpha/\lambda$.

Variance of the Hypothetical Means

The means μ in the population are distributed according to the Gamma distribution shown in (5.2.2) with parameters α and λ. The variance of the Gamma is α/λ^2. The μ's are the hypothetical means, so VHM = α/λ^2.

Bühlmann Credibility Parameter

Now we have the K parameter for Bühlmann Credibility in the Gamma-Poisson model: K = EPV/VHM = $(\alpha/\lambda)/(\alpha/\lambda^2) = \lambda$.

Calculating the Credibility Weighted Estimate

An insured is selected at random from a population of insureds whose average annual claims frequencies follow a Gamma distribution with parameters α and λ. Over the next Y years the insured is observed to have N claims. The credibility-weighted estimate of the average annual claims frequency for the insured is calculated as follows:

1. The observed annual claims frequency is N/Y.
2. The credibility of the Y years of observations is: $Z = Y/(Y+K) = Y/(Y+\lambda)$.
3. The prior hypothesis of the claims frequency is the mean of μ over the population: $E[\mu] = \alpha/\lambda$.
4. The credibility weighted estimate is:

$$\begin{aligned} &Z(N/Y) + (1-Z)(\alpha/\lambda) \\ &\qquad = \{Y/(Y+\lambda)\}(N/Y) + (1-\{Y/(Y+\lambda)\})(\alpha/\lambda) \\ &\qquad = \{N/(Y+\lambda)\} + \{\lambda/(Y+\lambda)\}(\alpha/\lambda) = \frac{\alpha+N}{\lambda+Y} \end{aligned} \tag{5.4.1}$$

This is exactly equal to the mean of the posterior Gamma distribution. **For the Gamma-Poisson, the estimates from using Bayesian Analysis and Bühlmann Credibility are equal.**[56]

Example 5.4.1: An insured is selected at random from a population whose average annual claims frequency follows a Gamma distribution with $\alpha = 3.0$ and $\lambda = 1.5$. (The insured's number of claims is Poisson distributed.) If the insured is observed to have 9 claims during the next three years, calculate the Bühlmann Credibility weighted estimate of the insured's average annual frequency.

[56]This is a special case of a general result for conjugate priors of members of "linear exponential families." This general result is beyond the scope of this chapter. In general, the estimates from Bayesian Analysis and Bühlmann Credibility may not be equal.

[*Solution:* This can be calculated directly from formula (5.4.1): $E[\mu \mid 9$ claims in 3 years$] = (3+9)/(1.5+3) = 2.67$. Alternatively if we want to go through all of the steps, EPV $= 3.0/1.5$, VHM $= 3.0/1.5^2 = 1.333$, $K = 2/1.333 = 1.5$, and $Z = 3/(3+1.5) = 2/3$. The prior estimate of the average annual frequency is $3.0/1.5 = 2$. The observed estimate is $9/3 = 3$. The credibility weighted estimate is $(2/3)(3) + (1/3)(2) = 2.67$.]

5.4. Exercises

5.4.1. An insured is selected at random from a population whose average annual claims frequency follows a Gamma distribution with $\alpha = 2.0$ and $\lambda = 8.0$. The distribution of the number of claims for each insured is Poisson. If the insured is observed to have 4 claims during the next four years, calculate the Bühlmann Credibility weighted estimate of the insured's average annual claims frequency.

5.4.2. You are given the following:

- A portfolio consists of 1,000 identical and independent risks.
- The number of claims for each risk follows a Poisson distribution with mean θ.
- Prior to the latest exposure period, θ is assumed to have a gamma distribution, with parameters $\alpha = 250$ and $\lambda = 2000$.

During the latest exposure period, the following loss experience is observed:

Number of Claims	Number of Risks
0	906
1	89
2	4
3	1
	1,000

Determine the mean of the posterior distribution of θ.

5.4.3. You are given the following:

- Number of claims follows a Poisson distribution with parameter μ.
- Prior to the first year of coverage, μ is assumed to have the Gamma distribution $f(\mu) = 1000^{150}\mu^{149}e^{-1000\mu}/\Gamma(150)$, $\mu > 0$.
- In the first year, 300 claims are observed on 1,500 exposures.
- In the second year, 525 claims are observed on 2,500 exposures.

After two years, what is the Bühlmann probability estimate of $E[\mu]$?

Use the following information for the next three questions:

Each insured has its accident frequency given by a Poisson Process with mean θ. For a portfolio of insureds, θ is distributed uniformly on the interval from 0 to 10.

5.4.4. What is the Expected Value of the Process Variance?

5.4.5. What is the Variance of the Hypothetical Means?

5.4.6. An individual insured from this portfolio is observed to have 7 accidents in a single year. Use Bühlmann Credibility to estimate the future accident frequency of that insured.

6. PRACTICAL ISSUES

This last section covers miscellaneous topics that are important in the application of credibility theory.

6.1. Examples of Choices for the Complement of Credibility

In the Bühlmann Credibility model, observations are made of a risk or group of risks selected from a population. Some characteristic of the risk, for example frequency, is to be estimated. If the observed frequency of the risk is not fully credible, then it is credibility weighted against the frequency estimate for the whole population.

This is the basic principle underlying experience rating plans that are widely used in calculating insurance premiums for commercial risks. The actual loss experience for an individual insured is used to adjust a "manual rate." (The "manual rate" is the rate from the insurance company's rate manual.) If the insured's experience has little or no credibility, then the insurance company will charge the insured the "manual rate" for the next time period. But, to the extent that the insured's prior experience is credible, the rate will be adjusted downward if the insured has better experience that the average risk to which the rate applies. Or, the rate will be adjusted upward if the insured has worse than average experience.

Credibility theory is used in a variety of situations. In practice, the actuary often uses judgment in the selection of a complement of credibility. The selected complement of credibility should be relatively stable and relevant to the random variable to be estimated.[57]

Example 6.1.1: The Richmond Rodents, your hometown's semi-professional baseball team, recruited a new pitcher, Roger Rocket, at the start of the season. In the first two games of a fifty game season, Roger had three hits in four at bats, a .750 batting average.

You have been asked to forecast Roger's batting average over the whole fifty game season. What complements of credibility would you consider using to make a credibility weighted forecast?

[*Solution:* Some complements of credibility you might consider include: Roger's batting average in high school, the total batting

[57] See Boor, "The Complement of Credibility," *PCAS* LXXXIII, 1996.

average of the Rodent's pitching staff over the last few years, or last season's batting average for all pitchers in the league.]

The following table lists a quantity to be estimated and a potential complement of credibility:

Value to be Estimated	Potential Complement of Credibility
New rate for Carpenters' workers compensation insurance	Last year's rate for Carpenters' workers compensation insurance[58]
Percentage increase for automobile rates in Wyoming	Weighted average of the inflation rates for automobile repair costs and medical care
Percentage change in rate for a rating territory in Virginia	Percent that the average statewide rate will change in Virginia
Rate to charge a specific Carpenter contracting business for workers compensation insurance	The average workers compensation insurance rate for Carpenters

In each case, the complement of credibility should be a reasonable estimate of the quantity of interest.

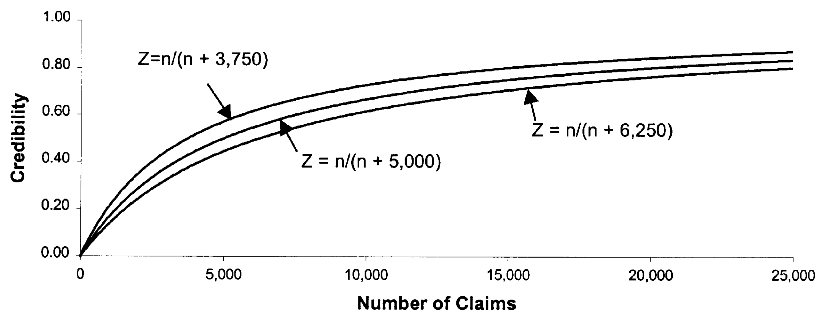
6.2. Impacts of Differences of Credibility Parameters

Relatively small differences of credibility formula parameters generally do not have big impacts on credibility estimates. In Graph 6 we show three credibility curves. The middle curve is a Bühlmann credibility formula $Z = n/(n + 5{,}000)$ where 5,000 was substituted for the K parameter. The bottom curve shows the results of increasing the K parameter by 25% to $K = 1.25 \times 5{,}000 = 6{,}250$. The top curve shows a smaller K value with $K = .75 \times 5{,}000 = 3{,}750$. Note that changing the Bühlmann credibility parameter by 25% has a relatively insignificant impact on the credibility.

[58] As with all complements of credibility, last year's rate may need to be adjusted for changes since then. For example, one might adjust it by the average rate change for all classes.

GRAPH 6

DIFFERENCE IN K PARAMETER



In practice, it is not necessary to precisely determine the best credibility formula parameters.[59] In fact, a good estimate is really the best that one can do.[60] Similar comments apply to Classical Credibility.

6.2. Exercises

6.2.1. The rate for workers compensation insurance for carpenters is currently \$18.00 (per \$100 of payroll.) The average rate over all classes is indicated to decrease 10%. The indicated rate for carpenters based on recent data is \$15.00. This recent data corresponds to \$3 million of expected losses. Calculate the proposed rate for carpenters in each following case:

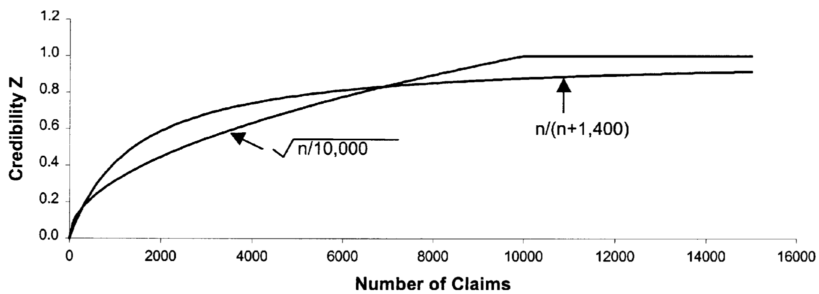
1. Using Bühlmann credibility parameter $K = \$5$ million.
2. Using Bühlmann credibility parameter $K = \$10$ million.
3. Using \$60 million of expected losses as the Standard for Full Credibility.

[59]Estimating K within a factor of two is usually sufficient. See Mahler's "An Actuarial Note on Credibility Parameters," *PCAS* LXXIII, 1986.

[60]Sometimes an actuary will judgmentally select K. Note that K is the volume of data that will be given 50% credibility by the Bühlmann credibility formula.

GRAPH 7

CLASSICAL VS. BÜHLMANN CREDIBILITY



6.3. Comparison of Classical and Bühlmann Credibility

Although the formulas look very different, Classical Credibility and Bühlmann Credibility can produce very similar results as seen in Graph 7.

The most significant difference between the two models is that Bühlmann Credibility never reaches $Z = 1.00$ which is an asymptote of the curve. Either model can be effective at improving the stability and accuracy of estimates.

Classical Credibility and Bühlmann Credibility formulas will produce approximately the same credibility weights if the full credibility standard for Classical Credibility, n_0, is about 7 to 8 times larger than the Bühlmann Credibility parameter K.[61]

Three estimation models were presented in this chapter: (1) Classical Credibility, (2) Bühlmann Credibility, and (3) Bayesian Analysis. For a particular application, the actuary can choose the model appropriate to the goals and data. If the goal is to generate

[61] This assumes that the Bühlmann Credibility formula is stated in terms of the number of claims. If it is not, then a conversion needs to be made. Suppose the Bühlmann Credibility formula is $E/(E+K)$ where E is earned premium. Further, assume that \$2,000 in earned premium is expected to generate one claim on average. Then, setting the full credibility standard n_0 equal to 7 or 8 times $K/2{,}000$ produces similar credibility weights.

the most accurate insurance rates, with least squares as the measure of fit, then Bühlmann Credibility may be the best choice. Bühlmann Credibility forms the basis of most experience rating plans. It is often used to calculate class rates in a classification plan. The use of Bühlmann Credibility requires an estimate of the EPV and VHM.

Classical Credibility might be used if estimates for the EPV and VHM are unknown or difficult to calculate. Classical Credibility is often used in the calculation of overall rate increases. Often it is simpler to work with. Bayesian Analysis may be an option if the actuary has a reasonable estimate of the prior distribution. However, Bayesian Analysis may be complicated to apply and the most difficult of the methods to explain to nonactuaries.

6.4. The Importance of Capping Results

Credibility is a linear process, and thus extreme cases can present difficulties requiring special attention. A properly chosen cap may not only add stability, but may even make the methodology more accurate by eliminating extremes. A class rating plan may have hundreds, if not thousands, of classifications. Credibility weighting can smooth out the fluctuations as rates or relativities are determined for each class, but with so many different classes, there will be extreme situations.

Suppose one is using a Classical Credibility model and that a full credibility standard has been selected so that the observations should be within 10% of the expected mean 95% of the time. This also means that 5% of the time the estimates will be more than 10% away from the expected mean. If there are 500 classes, then 25 classes on average will fall outside of the range, and some of these may be extreme.

In Classical Credibility, we assume the normal approximation. In practice this may not be a good assumption, particularly for situations where there is a limited volume of data. Insurance claim severities tend to be skewed to the right—in some cases highly skewed.[62] This is particularly true for some types of liability insurance and insurance providing fire coverage for highly valued property. A few large claims can result in very large pure premiums or high loss ratios. As the next section explains, this problem can be addressed by capping individual claims, but there is still the chance that several of these claims can occur during the same time period. Plus, even a capped loss can have a large impact if the volume of data is limited.

Credibility theory assumes that the underlying loss process is random and that events are independent. This may not always be true in the data you are working with. For example, weather events may produce a rash of claims. A spell of bitter cold can lead to many fire losses as people use less safe methods to heat their houses. Or, a couple of days of icy streets can produce many collision losses. The actuary can try to segregate out these special events, but it is not always possible to identify these and make appropriate adjustments.

Capping results is a good supplement to the credibility weighting process and makes the estimates more reliable. Users of the estimates may be more willing to accept them knowing that one or two events did not unduly affect the results.

6.4. Exercises

6.4.1. The workers compensation expected loss for crop dusting (via airplane) is \$20 per \$100 of payroll. Angel's Crop Dusting had \$300,000 of payroll and one claim for \$4 million over the experience period. Predict Angel's future expected losses per \$100 of payroll in each of the following cases:

1. You use the losses as reported and a Bühlmann credibility parameter of \$80,000 in expected losses.

[62] There is a significant probability that large claims will occur. The means for such distributions can be much larger than the medians.

2. You use the losses limited to \$200,000 per claim and a Bühlmann credibility parameter of \$50,000 in limited expected losses. Assume that limiting losses to \$200,000 per claim reduces them on average to 95% of unlimited losses.

6.5. Capping Data Used in Estimation

A technique commonly used in conjunction with credibility weighting is to cap large losses. Capping large losses can reduce the variance of observed incurred losses allowing more credibility to be assigned to the observations.

Suppose an actuary is calculating automobile insurance rates in a state. The rates will vary by geographic territory within the state. To limit the impact of individual large losses on territorial rates, each individual loss is capped at a selected amount, say \$100,000. If a loss is larger than \$100,000, then \$100,000 is substituted for the actual amount. So, only the first \$100,000 of any loss is included in a territory's loss data. The amounts above \$100,000 can be pooled and prorated across all territories in the state.

The Classical Credibility standard for full credibility is shown in formula (2.5.4):

$$n_F = n_0(1 + (\sigma_S/\mu_S)^2)$$

n_F is the expected number of claims required for a weight of 100% to be assigned to the data. Capping large losses reduces n_F by reducing the size of the coefficient of variation of the severity, $CV_S = \sigma_S/\mu_S$.

Example 6.5.1: Assume that claim losses can occur in three sizes in the following proportions:

Size of Loss	Proportion of Losses
\$1,000	80%
\$20,000	15%
\$100,000	5%

1. Calculate the coefficient of variation, $CV_S = \sigma_S/\mu_S$, for the severity of uncapped losses.

2. Calculate the coefficient of variation for severity if losses are capped at \$50,000.

[*Solution:* (1) Uncapped losses: $\mu_S = .80(1{,}000) + .15(20{,}000) + .05(100{,}000) = 8{,}800$. And, $\sigma_S = [.80(1{,}000)^2 + .15(20{,}000)^2 + .05(100{,}000)^2 - 8{,}800^2]^{1/2} = 21{,}985$. So, $CV_S = \sigma_S/\mu_S = 21{,}985/8{,}800 = 2.50$.

(2) Capped losses: $\mu_S = .80(1{,}000) + .15(20{,}000) + .05(50{,}000) = 6{,}300$. Note that 100,000 has been replaced by 50,000 in the calculation. And, $\sigma_S = [.80(1{,}000)^2 + .15(20{,}000)^2 + .05(50{,}000)^2 - 6{,}300^2]^{1/2} = 12{,}088$. So, $CV_S = \sigma_S/\mu_S = 12{,}088/6{,}300 = 1.92$.]

Using the results of example 6.5.1, the full credibility standard for uncapped losses is: $n_F = n_0(1 + (2.5)^2) = n_0(7.25)$. The full credibility standard capping losses at \$50,000 is: $n_F = n_0(1 + (1.92)^2) = n_0(4.69)$. Comparing these values we note that capping at \$50,000 reduced the number of claims required for full credibility by 35%.

Although Classical Credibility has been used as the setting to demonstrate capping, capping is also valuable when using a Bühlmann Credibility model. The process variance is reduced through capping which in turn usually lowers the K parameter. (Usually the Expected Value of the Process Variance is reduced more by capping than is the Variance of the Hypothetical Means.)

6.6. Estimating Credibility Parameters in Practice

This section discusses a few methods to estimate credibility parameters, but first it should be mentioned that judgment frequently plays a large role in the selection of credibility parameters.

The selection of credibility parameters requires a balancing of responsiveness versus stability. Larger credibility weights put more weight on the observations, which means that the current data has a larger impact on the estimate. The estimates are more responsive to current data. But, this comes at the expense of less stability in the estimates. Credibility parameters often are selected to reflect the actuary's desired balance between responsiveness and stability.

Classical Credibility

With Classical Credibility, P and k values must be chosen where P is the probability that observation X is within $\pm k$ percent of the mean μ. Bigger P's and smaller k's mean more stability but smaller credibility for the observations. A common choice for these values is $P = 90\%$ and $k = 5\%$, but there is no a priori reason why these choices are better than others.

When estimating the expected pure premium or loss ratio, the coefficient of variation of the severity distribution is needed to calculate the full credibility standard (see formula (2.5.4)). In practice this can be estimated from empirical data as demonstrated in the following example.

Example 6.6.1: A sample of 100 claims was distributed as follows:

Size of Claim	Number of Claims
\$1,000	85
\$5,000	10
\$10,000	3
\$25,000	2

Estimate the coefficient of variation of the claims severity based on this empirical distribution.

[*Solution:* The sample mean is: $\hat{\mu} = .85(1{,}000) + .10(5{,}000) + .03(10{,}000) + .02(25{,}000) = 2{,}150$. The sample standard deviation is: $\hat{\sigma} = \{[1/(100-1)][85(1{,}000-2{,}150)^2 + 10(5{,}000-2{,}150)^2 + 3(10{,}000-2{,}150)^2 + 2(25{,}000-2{,}150)^2]\}^{1/2} = 3{,}791$ where we are dividing by $(n-1)$ to calculate an unbiased estimate. The coefficient of variation is $CV_S = \hat{\sigma}/\hat{\mu} = 3{,}791/2{,}150 = 1.76$.]

With $P = 90\%$ and $k = 5\%$, the full credibility standard for frequency is $n_0 = 1082$ claims. The full credibility standard for the pure premium is $n_F = n_0(1 + CV_S^2) = 1{,}082(1 + 1.76^2) = 4{,}434$ with the coefficient of variation from example 6.6.1.

Bühlmann Credibility—Estimate EPV and VHM from Data

One way to estimate the K parameter in the formula $Z = N/(N+K)$ is to compute the numerator and denominator of K from empirical data.

Suppose that there are M risks in a population and that they are similar in size. Assume that we tracked the annual frequency year by year for Y years for each of the risks. The frequencies are:

$$\begin{bmatrix} X_{11} & X_{12} & \dots & X_{1Y} \\ X_{21} & X_{22} & \dots & X_{2Y} \\ \vdots & \vdots & & \vdots \\ X_{M1} & X_{M2} & \dots & X_{MY} \end{bmatrix}$$

Each row is a different risk and each column is a different year. So, X_{ij} represents the frequency for the i^{th} risk and j^{th} year.

Our goal is to estimate the expected frequency for any risk selected at random from the population. The credibility $Z = N/(N+K)$ will be assigned to the risk's observed frequency where N represents the number of years of observations. The complement of credibility will be assigned to the mean frequency for the population.

The following table lists estimators:

	Symbol	Estimator
Mean Frequency for Risk i	$\bar{X}_i$	$(1/Y)\sum_{j=1}^{Y} X_{ij}$
Mean Frequency for Population	$\bar{X}$	$(1/M)\sum_{i=1}^{M} \bar{X}_i$
Process Variance for Risk i	$\hat{\sigma}_i^2$	$[1/(Y-1)]\sum_{j=1}^{Y}(X_{ij}-\bar{X}_i)^2$
Expected Value of Process Variance	**EPV**	$(1/M)\sum_{i=1}^{M}\hat{\sigma}_i^2$
Variance of the Hypothetical Means	**VHM**	$[1/(M-1)]\sum_{i=1}^{M}(\bar{X}_i-\bar{X})^2 - EPV/Y$

The estimator for the Expected Value of the Process Variance is just an average of the usual estimators for each risk's process variance. The estimator for the Variance of the Hypothetical Means may not be intuitive, but we will use this result without a rigorous derivation.[63] After making the computations in the above table, we can set $K = \text{EPV}/\text{VHM}$. If the sample VHM is zero or negative, then the credibility can be assigned a value of 0.

The process variance of risk i, $\hat{\sigma}_i^2$, is the estimated process variance of the *annual* frequency for risk i. Since we have observed the frequency for Y years, we are able to estimate the variance in the annual frequency.

Example 6.6.2: There are two auto drivers in a particular rating class. The first driver was observed to have 2 claims in the first year, 0 in the second, 0 in the third, 1 in the fourth, and 0 in the fifth. The second driver had the following sequence of claims in years 1 through 5: 1, 1, 2, 0, 2. Estimate each of the values in the prior table.

[63] See Klugman, et al., *Loss Models: From Data to Decisions* for a more thorough and rigorous treatment of the estimation of Bühlmann credibility parameters.

[*Solution:* For driver #1, $\bar{X}_1 = (2+0+0+1+0)/5 = .6$ and $\hat{\sigma}_1^2 = [(2-.6)^2 + (0-.6)^2 + (0-.6)^2 + (1-.6)^2 + (0-.6)^2]/(5-1) = .80$. For driver #2, $\bar{X}_2 = (1+1+2+0+2)/5 = 1.2$ and $\hat{\sigma}_2^2 = [(1-1.2)^2 + (1-1.2)^2 + (2-1.2)^2 + (0-1.2)^2 + (2-1.2)^2]/(5-1) = .70$.

The population mean annual frequency is estimated to be $\bar{X} = (\bar{X}_1 + \bar{X}_2)/2 = (.6+1.2)/2 = .90$. The expected value of the process variance is $\text{EPV} = (\hat{\sigma}_1^2 + \hat{\sigma}_2^2)/2 = (.8+.7)/2 = .75$. The variance of the hypothetical means is $\text{VHM} = [(\bar{X}_1 - \bar{X})^2 + (\bar{X}_2 - \bar{X})^2]/(2-1) - \text{EPV}/5 = [(.6-.9)^2 + (1.2-.9)^2]/1 - .75/5 = .03$.]

The K parameter for the data in example 6.6.2 is $K = \text{EPV}/\text{VHM} = .75/.03 = 25$. The credibility that we would assign five years of experience is $Z = 5/(5+25) = 1/6$. Thus the estimated future claim frequency for the first driver is $(1/6)(.6) + (5/6)(.9) = .85$. Similarly, the estimated future claim frequency for the second driver is $(1/6)(1.2) + (5/6)(.9) = .95$. While in most practical applications there would be more than two drivers, this technique would apply in the same manner. When there are different sizes of insureds, for example commercial automobile fleets, the techniques are modified somewhat, but this is beyond the scope of this chapter.[64]

In practice there are many techniques used to estimate K. This was just one example of how to do so. It dealt with the simpler situation where every insured is of the same size.

Bühlmann Credibility—Estimate K from Best Fit to Data

Experience rating adjusts a policyholder's premium to reflect the policyholder's prior loss experience. If the policyholder has generated few insurance losses, then experience rating applies a credit to adjust the premium downward. And, if the policyholder has had worse than average loss experience, then debits increase

[64]See Klugman, et al., *Loss Models: From Data to Decisions*.

future premiums. A credibility formula of the form $Z = N/(N+K)$ is usually used to weight the policyholder's experience.[65]

One can estimate K by seeing which values of K would have worked well in the past. The goal is for each policyholder to have the same expected loss ratio after the application of experience rating. Let LR_i be the loss ratio for policyholder i where in the denominator we use the premiums after the application of experience rating.[66] Let LR_{AVE} be the average loss ratio for all policyholders. Then, define $D(K)$ to be:

$$D(K) = \sum_{\text{all } i} (LR_i - LR_{\text{AVE}})^2$$

The sum of the squares of the differences is a function of K, the credibility parameter that was used in the experience rating. The goal is to find a K that minimizes $D(K)$. This requires recomputing the premium that each policyholder would have been charged under a different K' value. This generates new LR_i's that are then put into the formula above and $D(K')$ is computed. Using techniques from numerical analysis, a $\hat{K}$ that minimizes $D(K)$ can be found.[67]

Another approach to calculating credibility parameters is linear regression analysis of a policyholder's current frequency, pure premium, etc. compared to prior experience. Suppose that we want to estimate next year's results based on the current year's. Then, using historical data for many policyholders we set up our regression equation:

$$\text{Observation in year } Y = m(\text{Observation in year } (Y-1)) + \text{Constant}$$

[65]The credibility weighting can be more complex and involve separate credibility factors for primary and excess losses.

[66]The experience modification relies on data from prior years, other than the one being tested, as well as the value of K.

[67]Criteria other than least squares can be used to see which credibilities would have worked well in the past.

The slope m from a least squares fit to the data turns out to be the Bühlmann credibility Z. The "Constant" term is $(1 - Z)$ (Overall Average).[68] After we have calculated the parameters in our model using historical data, then we can estimate future results using the model and recent data. Regression models can also be built using multiple years.

6.6. Exercises

6.6.1. Sue "Too Much Time on her Hands" Smith recorded her commute times to work in the morning while driving her husband's car during the week her sports car was in the shop. She also recorded her times for a week when she got her car back. Here were the results:

Trial	Husband's Car	Her Car
1	30 minutes	30 minutes
2	33 minutes	28 minutes
3	26 minutes	31 minutes
4	31 minutes	27 minutes
5	30 minutes	24 minutes

Using Bühlmann Credibility, Sue wants to estimate her expected commute time to work in her sports car. Calculate EPV, VHM, K, Z, and the credibility weighted estimate.

The next three problems share information and should be worked in order.

6.6.2. You observe the following experience for five insureds during years 1, 2, and 3 combined:

[68] See Mahler, "A Graphical Illustration of Experience Rating Credibilities," *PCAS* LXXXV, 1998 for more information.

Insured	Premiums (prior to experience mod)	Losses	Loss Ratio
1	1,000	600	60.0
2	500	200	40.0
3	2,000	1,100	55.0
4	1,500	700	46.7
5	3,000	2,200	73.3
Total	8,000	4,800	60.0

You will calculate experience modifications for these insureds using the formulas:

$$Z = P/(P+K), \qquad \text{and}$$
$$M = \{(L/P)Z + 60.0(1-Z)\}/60.0$$

where

Z = credibility, K = Bühlmann credibility parameter

P = premium, M = experience modification

L = losses, 60.0 = observed overall loss ratio.

What would the experience modifications be for each insured if you used $K = 1{,}000$?

6.6.3. You observe the following experience for these same five insureds during year five:

Insured	Premiums (prior to experience mod)	Losses	Loss Ratio
1	400	300	75.0
2	200	100	50.0
3	900	200	22.2
4	500	200	40.0
5	1,000	700	70.0
Total	3,000	1,500	50.0

Experience modifications are calculated using the data from years 1, 2, and 3 with various values of K.

For $K = 1{,}000$ what is the sum of the squared differences for year five between the modified loss ratios and the overall average loss ratio?

[Note: Modified premiums are calculated by multiplying premiums by modification factors. Modified loss ratios use modified premiums in the denominator.]

6.6.4. For what value of K would the sum of the squared differences in the previous problem be minimized? (Suggestion: use a computer to help you find the solution.)

REFERENCES

H = Historical, B = Basic, I = Intermediate, A = Advanced

H Bailey, Arthur L., "A Generalized Theory of Credibility," *Proceedings of the Casualty Actuarial Society*, 1945, 32:13–20.

H Bailey, Arthur L., "Credibility Procedures, Laplace's Generalization of Bayes' Rule and the Combination of Collateral Knowledge with Observed Data," *Proceedings of the Casualty Actuarial Society*, 1950, 37:7–23.

B Bailey, Robert A., and LeRoy J. Simon, "An Actuarial Note on the Credibility of Experience of a Single Private Passenger Car," *Proceedings of the Casualty Actuarial Society*, 1959, 46:159–164.

I Boor, Joseph A., "Credibility Based on Accuracy," *Proceedings of the Casualty Actuarial Society*, 1992, 79:166–185.

I Boor, Joseph A., "The Complement of Credibility," *Proceedings of the Casualty Actuarial Society*, 1996, 83:1–40.

I Brosius, Eric, "Loss Development Using Credibility." *CAS 2001 Exam 6 Web Notes*, 1993. http://www.casact.org/students/syllabus/2001/6web.htm (1 Oct. 1999).

H Bühlmann, Hans, "Experience Rating and Credibility," *ASTIN Bulletin*, 1967, 4, 3:200–207.

H Bühlmann, Hans and E. Straub, 1970. Glaubwürdigeit für schadensätze (Credibility for loss ratios), English translation, *ARCH*, 1972, 2.

I Conger, Robert F., "The Construction of Automobile Territories in Massachusetts," *Proceedings of the Casualty Actuarial Society*, 1987, 74:1–74.

B Dean, Curtis Gary, "An Introduction to Credibility," CAS *Forum*, Winter 1997, 55–66.

A De Vlyder, F. Etienne, *Advanced Risk Theory—A Self-Contained Introduction*, University of Brussels, Belgium, 1996.

H Dorweiler, Paul, "A Survey of Risk Credibility in Experience Rating," *Proceedings of the Casualty Actuarial Society*, 1971, 58:90–114. First published in *Proceedings of the Casualty Actuarial Society*, 1934, 21:1–25.

I Feldblum, Sholom, Discussion of "The Complement of Credibility," *Proceedings of the Casualty Actuarial Society*, 1998, 85:991–1033.

I Finger, Robert J., "Merit Rating for Doctor Professional Liability Insurance," *Proceedings of the Casualty Actuarial Society*, 1993, 80:291–352.

I Gerber, Hans, "A Teacher's Remark on Exact Credibility," *ASTIN Bulletin*, 1995, 25, 2:189–192.

I Gillam, William R., "Parametrizing the Workers Compensation Experience Rating Plan," *Proceedings of the Casualty Actuarial Society*, 1992, 79:21–56.

I Gillam, William R. and Richard H. Snader, "Fundamentals of Individual Risk Rating," 1992, available from the CAS.

I Goulet, Vincent, "On Approximations in Limited Fluctuation Credibility Theory," *Proceedings of the Casualty Actuarial Society*, 1997, 84:533–552.

I Goulet, Vincent, "Principles and Applications of Credibility," *Journal of Actuarial Practice*, 1998, 6:1, 2, 5–62.

I Hachemeister, Charles A., "Credibility for Regression Models with Application to Trend." In *Credibility, Theory and Applications, Proceedings of the Berkeley Actuarial Research Conference on Credibility*, 1975, 129–163, (reprinted with discussion by Al Quirin, in CAS *Forum*, Spring 1992, 307–348.)

I Heckman, Philip E., "Credibility and Solvency," *CAS Discussion Paper Program*, 1980, 116–152.

B Herzog, Thomas N., *An Introduction to Credibility*, Mad River Books, 1996.

B Hewitt, Charles C., "Credibility for Severity," *Proceedings of the Casualty Actuarial Society*, 1970, 57:148–171.

B Hewitt, Charles C., Discussion of "A Bayesian View of Credibility," *Proceedings of the Casualty Actuarial Society*, 1965, 52:121–127.

B Hossack, I. B., J. H. Pollard, and B. Zehnwirth, *Introductory Statistics with Applications in General Insurance*, Cambridge University Press, New York, 1983.

I Insurance Services Office, *Report of the Credibility Subcommittee: Development and Testing of Empirical Bayes Credibility Procedures for Classification Ratemaking*, ISO, New York, 1980.

A Jewell, W. S., "Credible Means Are Exact Bayesian for Exponential Families," *ASTIN Bulletin*, 1974, 8, 1:77–90.

A Jewell, W. S., "The Use of Collateral Data in Credibility Theory: A Hierarchical Model," *Giornale dell'Istituto Italiano degli Attuari*, 1975, 38:1–16.

A Klugman, Stuart A., "Credibility for Classification Ratemaking Via the Hierarchical Normal Linear Model," *Proceedings of the Casualty Actuarial Society*, 1987, 74:272–321.

I Klugman, Stuart A., Harry H. Panjer, and Gordon E. Willmot, *Loss Models: From Data to Decisions*, John Wiley and Sons, New York, 1998.

B Longley-Cook, Lawrence H., "An Introduction to Credibility Theory," *Proceedings of the Casualty Actuarial Society*, 1962, 49:194–221.

B Mahler, Howard C., "A Graphical Illustration of Experience Rating Credibilities, *Proceedings of the Casualty Actuarial Society*, 1998, 85:654–688.

A Mahler, Howard C., "A Markov Chain Model for Shifting Risk Parameters," *Proceedings of the Casualty Actuarial Society*, 1997, 84:581–659.

B Mahler, Howard C., "An Actuarial Note on Credibility Parameters," *Proceedings of the Casualty Actuarial Society*, 1986, 73:1–26.

I Mahler, Howard C., "An Example of Credibility and Shifting Risk Parameters," *Proceedings of the Casualty Actuarial Society*, 1990, 77:225–308.

B Mahler, Howard C., "Credibility: Practical Applications," *CAS Forum*, Fall 1989, 187–199.

A Mahler, Howard C., "Credibility With Shifting Risk Parameters, Risk Heterogeneity and Parameter Uncertainty," *Proceedings of the Casualty Actuarial Society*, 1998, 85:455–653.

A Mahler, Howard C., Discussion of "An Analysis of Experience Rating," *Proceedings of the Casualty Actuarial Society*, 1987, 74:119–189.

I Mahler, Howard C., Discussion of "Parametrizing the Workers Compensation Experience Rating Plan," *Proceedings of the Casualty Actuarial Society*, 1993, 80:148–183.

B Mahler, Howard C., "Introduction to Basic Credibility," *CAS Forum*, Winter 1997, 67–103.

A Mahler, Howard C., "Workers' Compensation Classification Relativities," *CAS Forum*, Fall 1999, 425–462.

B Mayerson, Allen L., "A Bayesian View of Credibility," *Proceedings of the Casualty Actuarial Society*, 1964, 51:85–104.

B Mayerson, Allen L., Donald A. Jones, and Newton Bowers Jr., "The Credibility of the Pure Premium," *Proceedings of the Casualty Actuarial Society*, 1968, 55:175–185.

A Meyers, Glenn G., "An Analysis of Experience Rating," *Proceedings of the Casualty Actuarial Society*, 1985, 72:278–317.

A Meyers, Glenn G., "Empirical Bayesian Credibility for Workers Compensation Classification Ratemaking," *Proceedings of the Casualty Actuarial Society*, 1984, 71:96–121.

H Michelbacher, Gustav F., "The Practice of Experience Rating," *Proceedings of the Casualty Actuarial Society*, 1918, 4:293–324.

H Mowbray, A. H., "How Extensive a Payroll Exposure is Necessary to Give a Dependable Pure Premium?" *Proceedings of the Casualty Actuarial Society*, 1914, 1:25–30.

H Perryman, Francis S., "Experience Rating Plan Credibilities," *Proceedings of the Casualty Actuarial Society*, 1971, 58:143–207. First published in *Proceedings of the Casualty Actuarial Society*, 1937, 24:60–125.

H Perryman, Francis S., "Some Notes on Credibility," *Proceedings of the Casualty Actuarial Society*, 1932, 19:65–84.

B Philbrick, Stephen W., "An Examination of Credibility Concepts," *Proceedings of the Casualty Actuarial Society*, 1981, 68:195–219.

I Robbin, Ira, "A Bayesian Credibility Formula for IBNR Counts," *Proceedings of the Casualty Actuarial Society*, 1986, 73:129–164.

H Uhthoff, Dunbar R., "The Compensation Experience Rating Plan—A Current Review," *Proceedings of the Casualty Actuarial Society*, 1959, 46:285–299.

I Van Slyke, E. Oakley, "Credibility-Weighted Trend Factors," *Proceedings of the Casualty Actuarial Society*, 1981, 68:160–171.

B Van Slyke, E. Oakley, "Credibility for Hiawatha," *CAS Forum*, Fall 1995, 281–298.

I Venter, Gary G., "Classical Partial Credibility with Application to Trend," *Proceedings of the Casualty Actuarial Society*, 1986, 73:27–51.

I Venter, Gary G. Credibility. Chap. 7 in *Foundations of Casualty Actuarial Science*. First Edition. New York: Casualty Actuarial Society, 1990.

A Venter, Gary G., "Structured Credibility in Applications—Hierarchical, Multi-Dimensional and Multivariate Models," *ARCH*, 1985, 2.

I Waters, H. R., "An Introduction to Credibility Theory," Institute of Actuaries Study Note.

H Whitney, Albert W., "The Theory of Experience Rating," *Proceedings of the Casualty Actuarial Society*, 1918, 4:274–292.

APPENDIX

FREQUENCY AND LOSS DISTRIBUTIONS

Actuaries commonly use distributions to model the number of claims and sizes of claims. This appendix will give key facts about the most commonly used frequency distributions and some of the more commonly used loss distributions.

Frequency Distributions

Binomial Distribution

Support: $x = 0,1,2,3\ldots,n$ Parameters: $1 > p > 0$, $n \geq 1$. Let $q = 1 - p$.

Probability density function: $f(x) = \binom{n}{x} p^x q^{n-x}$

Mean = np

Variance = npq

Special Case: For $n = 1$ one has a Bernoulli Distribution

Poisson Distribution

Support: $x = 0,1,2,3\ldots$ Parameters: $\lambda > 0$

Probability density function: $f(x) = \lambda^x e^{-\lambda}/x!$

Mean = λ

Variance = λ

Negative Binomial Distribution

Support: $x = 0,1,2,3\ldots$ Parameters: $k \geq 0$, $0 < p < 1$. Let $q = 1 - p$

Probability density function: $f(x) = \binom{x+k-1}{x} p^k q^x$

Mean = $kq/p = k(1-p)/p$

Variance = $kq/p^2 = k(1-p)/p^2$

Special Case: For $k = 1$ one has a Geometric Distribution

Loss Distributions

Exponential Distribution

Support: $x > 0$ Parameters: $\lambda > 0$

Distribution Function: $F(x) = 1 - e^{-\lambda x}$

Probability density function: $f(x) = \lambda e^{-\lambda x}$

Moments: $E[X^n] = (n!)/\lambda^n$

Mean = $1/\lambda$

Variance = $1/\lambda^2$

Gamma Distribution

Support: $x > 0$ Parameters: $\alpha, \lambda > 0$

Distribution Function: $F(x) = \Gamma(\alpha; \lambda x)$

Probability density function: $f(x) = \lambda^\alpha x^{\alpha-1} e^{-\lambda x}/\Gamma(\alpha)$

Moments: $E[X^n] = \prod_{i=0}^{n-1} (\alpha + i)/\lambda^n = \lambda^{-n}\Gamma(\alpha+n)/\Gamma(\alpha)$

Mean = α/λ

Variance = α/λ^2

Special Case: For $\alpha = 1$ one has an Exponential Distribution

Weibull Distribution

Support: $x > 0$ Parameters: $c, \tau > 0$

Distribution Function: $F(x) = 1 - \exp(-cx^\tau)$

Probability density function: $f(x) = c\tau x^{\tau-1}\exp(-cx^\tau)$

Moments: $E[X^n] = \Gamma(1 + n/\tau)/c^{n/\tau}$

Special Case: For $\tau = 1$ one has an Exponential Distribution

LogNormal Distribution

Support: $x > 0$

Parameters: $-\infty < \mu < +\infty,\ \sigma > 0$

Distribution Function: $F(x) = \Phi[\ln(x) - \mu/\sigma]$

Probability density function: $f(x) = \exp[-.5(\{\ln(x) - \mu\}/\sigma)^2]/\{x\sigma\sqrt{2\pi}\}$

Moments: $E[X^n] = \exp[n\mu + .5n^2\sigma^2]$

Mean $= \exp(\mu + .5\sigma^2)$

Variance $= \exp(2\mu + \sigma^2)\{\exp(\sigma^2) - 1\}$

Pareto Distribution

Support: $x > 0$

Parameters: $\alpha,\ \lambda > 0$

Distribution Function: $F(x) = 1 - (\lambda/(\lambda + x))^\alpha = 1 - (1 + x/\lambda)^{-\alpha}$

Probability density function: $f(x) = (\alpha\lambda^\alpha)(\lambda + x)^{-(\alpha+1)} = (\alpha/\lambda)(1 + x/\lambda)^{-(\alpha+1)}$

Moments: $E[X^n] = \lambda^n n!/\prod_{i=1}^{n}(\alpha - i) \qquad \alpha > n$

Mean $= \lambda/(\alpha - 1) \qquad \alpha > 1$

Variance $= \lambda^2\alpha/\{(\alpha - 2)(\alpha - 1)^2\} \qquad \alpha > 2$

SOLUTIONS

Solution 2.2.1: $\Phi(2.576) = .995$, so that $y = 2.576$. $n_0 = y^2/k^2 = (2.576/.025)^2 =$ **10,617**.

Solution 2.2.2: $\Phi(2.326) = .99 = (1 + .98)/2$, so that $y = 2.326$. $n_0 = y^2/k^2 = (2.326/.075)^2 =$ **962**.

Solution 2.2.3: $n_0 = y^2/k^2$. Therefore, $y = k\sqrt{n_0} = .06\sqrt{900} = 1.80$. $P = 2\Phi(y) - 1 = 2\Phi(1.80) - 1 = (2)(.9641) - 1 =$ **.9282**.

Solution 2.2.4: For Y risks the mean is $.05Y$ and variance is $.09Y$. (The means and variances of independent variables each add.) Thus, the $\pm 2\%$ error bars correspond to $\pm(.02)(.05Y)$. The standard deviation is $.3(Y^{.5})$. "94% of the time," corresponds to 1.881 standard deviations, since $\Phi(1.881) = 97\%$. Thus, we set the error bars equal to 1.881 standard deviations: $(1.881)(.3)Y^{.5} = (.02)(.05Y)$. Therefore, $Y = (1.881/.02)^2(.09)/.05^2 =$ **318,434**.

Comment: In terms of claims instead of exposures the full credibility standard would be $318{,}434 \times .05 = 15{,}922 = 8{,}845 \times 1.8 =$ 8,845 (variance/mean). If the claim frequency were Poisson, the variance equals the mean and one would get a standard for full credibility of 8,845 claims; in this case since the Poisson assumption does not hold one must multiply by an additional term of $(.09/.05)$ = variance/mean. Since the variance is larger than the mean, we need more claims to limit the fluctuations than would be the case with a Poisson frequency.

Solution 2.2.5: Let x be the number of respondents and let p be the true percentage of yes respondents in the total population. The result of the poll is a Binomial Distribution with variance $xp(1-p)$. Thus the variance of the average is $(1/x^2)$ times this or $p(1-p)/x$. Using the Normal Approximation, 95% probability corresponds to ± 1.96 standard deviations of the mean of p. Thus we want $(.07)(p) = (1.96)\sqrt{(p(1-p)/x)}$. $\sqrt{x} = (1.96)(\sqrt{(1-p)/p})/.07$. $x = 784((1/p) - 1)$. As p gets smaller x approaches infinity. However, we assume $p \geq .2$ so that $x \leq 784(5 - 1) =$ **3,136**.

Comment: The 3,136 respondents are similar to 3,136 exposures. If one has at least a 20% chance of a yes response, then the expected number of yeses is at least (3,136)(.2) = 627. This is similar in concept to 627 expected claims. The general form of the standard for full credibility is in terms of expected claims: $(\sigma_f^2/\mu_f)(y^2/k^2)$. In this case, $k = .07$, $P = 95\%$ and $y = 1.960$. $\sigma_f^2/\mu_f = npq/(np) = q$. Thus the standard for full credibility in terms of expected claims would be: $q(1.960/.07)^2 = 784q$. In terms of exposures it would be: $784q/p = 784(1/p - 1)$. For p between .2 and .8, this expression is maximized when $p = .2$ and is then $784(5 - 1) = 3,136$ exposures.

Solution 2.2.6: For frequency, the general formula for the Standard for Full Credibility is: $(\sigma_f^2/\mu_f)\{y^2/k^2\}$. Assuming y and k (not the parameter of a Negative Binomial, but rather the tolerance around the true mean frequency) are fixed, then the Standard for Full Credibility is proportional to the ratio of the variance to the mean. For the Poisson this ratio is one. For a Negative Binomial with parameters p and k, this ratio is: $(kq/p^2)/(kq/p) = 1/p$. Thus the second Standard is $1/p = 1/.7 = 1.429$ times the first standard.

Comment: The Negative Binomial has a larger variance than the Poisson, so there is more random fluctuation, and therefore the standard for Full Credibility is larger. For the Poisson $\sigma_f^2/\mu_f = 1$. For the Negative Binomial the variance is greater than the mean, so $\sigma_f^2/\mu_f > 1$. Thus for the Negative Binomial the Standard for Full Credibility is larger than the Poisson case, all else equal.

Solution 2.2.7: Since the full credibility standard is inversely proportional to the square of k: $n_0 = y^2/k^2$, $X/Y = (10\%/5\%)^2 = 4$. Alternately, one can compute the values of X and Y assuming one is dealing with the standard for Frequency. For $k = 5\%$ and $P = 90\%$: $\Phi(1.645) = .95 = (1 + .90)/2$, so that $y = 1.645$, $n_0 = y^2/k^2 = (1.645/.05)^2 = 1{,}082 = X$. For $k = 10\%$ and $P = 90\%$:

$\Phi(1.645) = .95 = (1 + .90)/2$, so that $y = 1.645$, $n_0 = y^2/k^2 = (1.645/.10)^2 = 271 = Y$. Thus $X/Y = 1{,}082/271 = \mathbf{4}$.

Comment: As the requirement gets less strict, for example k=10% rather than 5%, the number of claims needed for Full Credibility decreases. If one is dealing with the standard for pure premiums rather than frequency, then both X and Y have an extra factor of $(1 + CV^2)$, which doesn't effect X/Y.

Solution 2.3.1: $y = 1.645$ since $\Phi(1.645) = .95$. $n_0 = (y/k)^2 = (1.645/.01)^2 = 27060$. For severity, the Standard For Full Credibility is: $n_0 CV^2 = (27{,}060)(6{,}000{,}000/1{,}000^2) = (27{,}060)(6) =$ **162,360**.

Solution 2.3.2: $n_0 CV_S^2 =$ Full Credibility Standard for Severity. n_0 is the same for both risks A and B. Since the means of the severity of A and B are the same, but the standard deviation of severity for B is twice as large as A's, then $CV_S^2(\mathrm{B}) = 4CV_S^2(\mathrm{A})$. This implies that $n_0 CV_S(\mathrm{B}) = 4n_0 CV_S(\mathrm{A}) = \boldsymbol{4N}$.

Solution 2.4.1: $\sigma_{PP}^2 = \mu_F \sigma_S^2 + \mu_S^2 \sigma_F^2 = (13)(200{,}000) + (300)^2(37)$ $=$ **5,930,000**.

Solution 2.4.2: Frequency is Bernoulli with $p = 2/3$, with mean $= 2/3$ and variance $= (2/3)(1/3) = 2/9$. Mean severity = 7.1, variance of severity $= 72.1 - 7.1^2 = 21.69$. Thus, $\sigma_{PP}^2 = \mu_F \sigma_S^2 + \mu_S^2 \sigma_F^2 = (2/3)(21.69) + (7.1^2)(2/9) =$ **25.66**.

For the severity the mean and the variance are computed as follows:

Probability	Size of Claim	Square of Size of Claim
20%	2	4
50%	5	25
30%	14	196
Mean	7.1	72.1

Solution 2.4.3: The average Pure Premium is 106. The second moment of the Pure Premium is 16,940. Therefore, the variance = 16,940 – 106^2 = **5,704**.

Situation	Probability	Pure Premium	Square of P.P.
1 claim @ 50	60.0%	50	2,500
1 claim @ 200	20.0%	200	40,000
2 claims @ 50 each	7.2%	100	10,000
2 claims: 1 @ 50 & 1 @ 150	9.6%	200	40,000
2 claims @ 150 each	3.2%	300	90,000
Overall	100.0%	106	16,940

For example, the chance of 2 claims with one of size 50 and one of size 150 is the chance of having two claims times the chance given two claims that one will be 50 and the other 150 = (.2)(2)(.6)(.4) = 9.6%. In that case the pure premium is 50 + 150 = 200. One takes the weighted average over all the possibilities.

Comment: Note that the frequency and severity are *not* independent.

Solution 2.4.4: Since the frequency and severity are independent, the process variance of the Pure Premium = (mean frequency)(variance of severity) + (mean severity)^2 (variance of frequency) = .25 [(variance of severity) + (mean severity)^2] = .25 (2nd moment of the severity)

$$= (.25/5{,}000)\int_0^{5000} x^2 dx$$

$$= (.25/5{,}000)(5{,}000)^3/3 = \mathbf{2{,}083{,}333}.$$

Solution 2.4.5: The mean severity = $\exp(\mu + .5\sigma^2)$ = exp(4.32) = 75.19. Thus the mean aggregate losses is (8,200)(75.19) = **616,547**. The second moment of the severity = $\exp(2\mu + 2\sigma^2)$ = exp(9.28) = 10,721. Thus since the frequency is Poisson and independent of the severity the variance of the aggregate losses

$= \mu_F$ (2nd moment of the severity) = (8,200)(10,721) = **87.91 million**.

Solution 2.4.6: For a Poisson frequency, with independent frequency and severity, the variance of the aggregate losses = μ_F (2nd moment of the severity) = (0.5)(1,000) = 500.

Solution 2.4.7: Since we have a Poisson Frequency, the Process Variance for each type of claim is given by the mean frequency times the second moment of the severity. For example, for Claim Type Z, the second moment of the severity is $(1{,}500^2 + 2{,}000{,}000) = 4{,}250{,}000$. Thus, for Claim Type Z the process variance of the pure premium is: (.01)(4,250,000) = 42,500. Then the process variances for each type of claim add to get the total variance, 103,570.

Type of Claim	Mean Frequency	Mean Severity	Square of Mean Severity	Variance of Severity	Process Variance of P.P.
W	0.02	200	40,000	2,500	850
X	0.03	1,000	1,000,000	1,000,000	60,000
Y	0.04	100	10,000	0	400
Z	0.01	1,500	2,250,000	2,000,000	42,500
Sum					**103,750**

Solution 2.5.1: For a LogNormal, the mean severity $= \exp(\mu + .5\sigma^2) = \exp(4.32) = 75.19$. The second moment of the severity $= \exp(2\mu + 2\sigma^2) = \exp(9.28) = 10{,}721$. Thus $1 + CV^2$ = second moment divided by square of the mean $= 10{,}721/75.19^2 = 1.896$. (Note that for the LogNormal Distribution: $1 + CV^2 = \exp(\sigma^2) = \exp(.8^2) = 1.8965$.) $y = 1.645$ since $\Phi(1.645) = .95 = (1 + .90)/2$. Therefore, $n_0 = y^2/k^2 = (1.645/.025)^2 = 4{,}330$. Therefore, $n_F = n_0(1 + CV^2) = (4{,}330)(1.896) =$ **8,210 claims**.

Solution 2.5.2: Square of Coefficient of Variation = (2 million)/$(1{,}000^2) = 2$. The Normal distribution has a 99.5% chance of being less than 2.575. Thus $y = 2.575$. $k = 10\%$. Therefore, in terms of number of claims the full credibility standard is = $y^2(1+CV^2)/k^2 = (2.575^2)(1+2)/10\%^2 = 1989$ claims. This is equivalent to $1989/.04 =$ **49,725 policies**.

Solution 2.5.3: The severity has a mean of 16.67 and a second moment of 416.67:

$$\int_0^{50} xf(x)dx = .0008\int_0^{50}(50x - x^2)dx$$

$$= .0008(25x^2 - x^3/3)\Big]_0^{50} = 16.67$$

$$\int_0^{50} x^2 f(x)dx = .0008\int_0^{50}(50x^2 - x^3)dx$$

$$= .0008(50x^3/3 - x^4/4)\Big]_0^{50} = 416.67$$

$1 + CV^2 = E[X^2]/E^2[X] = 416.67/16.67^2 = 1.5.$

The standard for Full Credibility for the pure premiums for $k = 2.5\%$ is, therefore, $n_F = n_0(1+CV^2) = (5{,}000)(1.5) = 7{,}500$. For k=9% we need to multiply by $(2.5/9)^2$ since the full credibility standard is inversely proportional to k^2. $7{,}500(2.5/9)^2 =$ **579**.

Solution 2.5.4: We have $y = 2.576$ since $\Phi(2.576) = .995$. Therefore, $n_0 = (y/k)^2 = (2.576/.10)^2 = 663$. $n_F = n_0(1+CV^2)$, therefore, $CV = \sqrt{(n_F/n_0) - 1} = \sqrt{(2{,}000/663) - 1} =$ **1.42**.

Solution 2.5.5: We have $y = 1.960$ since $\Phi(1.960) = .975$. Therefore, $n_0 = (y/k)^2 = (1.960/.20)^2 = 96$. The mean severity is $(10)(.5) + (20)(.3) + (50)(.2) = 21$. The variance of the severity is: $(11^2)(.5) + (1^2)(.3) + (29^2)(.2) = 229$. Thus, the coefficient of variation squared $= 229/21^2 = .519$. $n_F = n_0(1+CV^2) = 96(1.519) =$ **146**.

Solution 2.5.6: The Poisson assumption does not apply so we use formula [2.5.5]: $n_F = (y^2/k^2)(\sigma_f^2/\mu_f + \sigma_s^2/\mu_s^2)$. We have $y =$

1.960 since $\Phi(1.960) = .975$. Therefore, $(y/k)^2 = (1.960/.20)^2 = 96$. The mean severity is $(10)(.5) + (20)(.3) + (50)(.2) = 21$. The variance of the severity is: $(11^2)(.5) + (1^2)(.3) + (29^2)(.2) = 229$. Therefore, $\sigma_s^2/\mu_s^2 = 229/21^2 = .519$. The variance of the frequency is twice the mean, so $\sigma_f^2/\mu_f = 2$. The answer is: $n_F = (96)(2 + .519) = \mathbf{242}$. Because of the greater variance in frequency relative to the mean, more claims are required for the Negative Binomial frequency than the Poisson frequency in the previous problem.

Solution 2.5.7: $\Phi(2.327) = .99$, so $y = 2.327$. For frequency, the standard for full credibility is $(2.327/.025)^2 = 8{,}664$. On the other hand, the Standard for Full Credibility for the pure premium: $\Phi(1.645) = .95$, so $y = 1.645$. Thus, $8{,}664 = n_F = (y^2/k^2)(1 + CV^2) = (1.645^2/k^2)(1 + 3.5^2) = 35.85/k^2$. Thus, $k = \sqrt{(35.85/8{,}664)} = \mathbf{.064}$.

Solution 2.5.8: The mean of the severity distribution is 100,000 /2 = 50,000. The Second Moment of the Severity Distribution is the integral from 0 to 100,000 of $x^2 f(x)$, which is $100{,}000^3/3(100{,}000)$. Thus, the variance is $100{,}000^2/3 - 50{,}000^2 = 833{,}333{,}333$. Thus, the square of the coefficient of variation is $833{,}333{,}333/50{,}000^2 = 1/3$. $k = 5\%$ (within $\pm 5\%$) and since $P = .95$, $y = 1.960$ since $\Phi(1.960) = (1 + P)/2 = .975$.

The Standard for Full Credibility Pure Premium $= (y/k)^2 (1 + CV^2) = (1.96/.05)^2\ (1 + 1/3) = 1537\ (4/3) =$ **2,049 claims**.

Solution 2.5.9: $n_F = (y/k)^2(1 + CV^2)$. If the CV goes from 2 to 4, and k doubles then the Standard for Full Credibility is multiplied by $\{(1 + 4^2)/(1 + 2^2)\}/2^2 = (17/5)/4$. Thus, the Standard for Full Credibility is altered to: $(1{,}200)(17/5)/4 = \mathbf{1{,}020}$.

Solution 2.6.1: $Z = \sqrt{(300/2{,}000)} = \mathbf{38.7\%}$.

Solution 2.6.2: Since the credibility is proportional to the square root of the number of claims, we get $(36\%)(\sqrt{10}) = 114\%$. However, the credibility is limited to **100%**.

Solution 2.6.3: $Z = \sqrt{(803/2{,}500)} = .567$. Observed average cost per claim is $9{,}771{,}000/803 = 12{,}168$. Thus the estimated severity $= (.567)(12{,}168) + (1 - .567)(10{,}300) =$ **\$11,359**.

Solution 2.6.4: The expected number of claims is $(.06)(2{,}000) = 120$. $Z = \sqrt{(120/3{,}300)} = 19.1\%$.

Comment: The Standard for Full Credibility could be given in terms of house-years rather than claims: for $3{,}300/.06 = 55{,}000$ house-years one expects 3,300 claims.

The credibility for 2,000 house years is: $\sqrt{(2{,}000/55{,}000)} = 19.1\%$.

Solution 2.6.5: $\Phi(2.17) = .985$, so that $y = 2.17$. $n_0 = y^2/k^2 = (2.17/.04)^2 = 2{,}943$. For the Gamma Distribution, the mean is α/λ, while the variance is α/λ^2. Thus the coefficient of variation is $(\text{variance}^{.5})/\text{mean} = \{\alpha/\lambda^2\}^{.5}/\alpha/\lambda = 1/\alpha^{.5}$. So for the given Gamma with $\alpha = 1.5$: $(1 + CV^2) = 1 + 1/1.5 = 1.667$. $n_F = n_0(1 + CV^2) = (2{,}943)(1.667) = 4{,}905$. $Z = \sqrt{(150/4{,}905)} = 17.5\%$.

Solution 2.6.6: The credibility $Z = \sqrt{(600/5{,}400)} = 1/3$. Thus the new estimate is: $(1/3)(1{,}200) + (1 - 1/3)(1{,}000) =$ **\$1,067**.

Solution 2.6.7: $(1 + P)/2 = (1.96)/2 = .98$. Thus, $y = 2.054$ since $\Phi(2.054) = .98$. The standard for full credibility is: $(y^2/k^2)(1 + CV^2) = (2.054/.10)^2 \ (1 + .6^2) = 574$ claims. Thus, we assign credibility of $Z = \sqrt{(213/574)} =$ **60.9%**.

Solution 2.6.8: $k = .05$ and $P = .90$. $y = 1.645$, since $\Phi(1.645) = .95 = (1 + P)/2$. $n_0 = y^2/k^2 = (1.645/.05)^2 = 1{,}082$. The mean of the severity distribution is 100,000. The second moment of the severity is the integral of $x^2/200{,}000$ from 0 to 200,000, which is $200{,}000^2/3$. Thus, the variance is 3,333,333,333. The square of the coefficient of variation is variance/mean^2 = $3{,}333{,}333{,}333/100{,}000^2 = .3333$. Thus, $n_F = n_0(1 + CV^2) = (1{,}082)(1.333) = 1{,}443$. For 1,082 claims $Z = \sqrt{(1{,}082/1{,}443)} = \sqrt{(3/4)} =$ **.866**.

Solution 2.6.9: $Z = \sqrt{(10{,}000/17{,}500)} = 75.6\%$. Thus, the new estimate $= (25 \text{ million})(.756) + (20 \text{ million})(1 - .756) =$ **\$23.78 million**.

Solution 2.6.10: $Z = \sqrt{(n/\text{standard for full credibility})} = \sqrt{(n/2{,}000)}$. Setting the credibility equal to .6: $.6 = \sqrt{(n/2{,}000)}$. Therefore, $n = (.6^2)(2{,}000) =$ **720 claims**.

Solution 2.6.11: We are given $k = 5\%$ and $P = 90\%$, therefore, we have $y = 1.645$ since $\Phi(1.645) = .95 = (1 + P)/2$. Therefore, $n_0 = (y/k)^2 = (1.645/.05)^2 = 1{,}082$. The partial credibility is given by the square root rule: $Z = \sqrt{(500/1{,}082)} =$ **.68**.

Solution 3.1.1:

Type of Risk	A Priori Probability	Process Variance
A	0.60	0.16
B	0.25	0.21
C	0.15	0.24
Average		**0.1845**

For a Bernoulli the process variance is pq. For example, for Risk Type B, the process variance = (.3)(.7) = .21.

Solution 3.1.2:

Type of Risk	A Priori Probability	Mean	Square of Mean
A	0.60	0.2	0.04
B	0.25	0.3	0.09
C	0.15	0.4	0.16
Average		0.2550	0.0705

Thus, the Variance of the Hypothetical Means $= .0705 - .255^2 =$ **.0055**.

Solution 3.1.3:

Number of Claims	Probability for Youthful	n	Square of n
0	0.80	0	0
1	0.15	1	1
2	0.04	2	4
3	0.01	3	9
Average		0.2600	0.4000

Thus, the process variance for the Youthful drivers is $.4 - .26^2 = .3324$.

Number Claims	Probability for Adult	n	Square of n
0	0.90	0	0
1	0.08	1	1
2	0.02	2	4
3	0.00	3	9
Average		0.1200	0.1600

Thus, the process variance for the Adult drivers is $.16 - .12^2 = .1456$.

Thus, the Expected Value of the Process Variance = (.1456)(91%) + (.3324)(9%) = **.162**.

Solution 3.1.4:

Type of Driver	A Priori Probability	Mean	Square of Mean
Youthful	0.0900	0.2600	0.06760
Adult	0.9100	0.1200	0.01440
Average		0.1326	0.01919

Thus, the Variance of the Hypothetical Means $= .01919 - .1326^2 = \mathbf{.00161}$.

Solution 3.1.5:

A Number of Claims	B A Priori Probability	C Col. A Times Col. B.	D Square of Col. A. Times Col. B
0	0.55000	0.00000	0.00000
1	0.30000	0.30000	0.30000
2	0.10000	0.20000	0.40000
3	0.04000	0.12000	0.36000
4	0.01000	0.04000	0.16000
Sum	1	0.6600	1.22000

Each insured's frequency process is given by a Poisson with parameter θ, with θ varying over the group of insureds. Then the process variance for each insured is θ. Thus the expected value of the process variance is estimated as follows:

$$E_\theta[\mathrm{Var}[X \mid \theta]] = E_\theta[\theta] = \text{overall mean} = \mathbf{.66}.$$

Solution 3.1.6: Consulting the table in the prior solution, the mean number of claims is .66. The second moment of the distribution of the number of claims is 1.22. So, the Total Variance for the distribution of the number of claims is: $1.220 - .660^2 = \mathbf{.784}$.

Solution 3.1.7: Using the solutions to the previous questions, we estimate the Variance of the Hypothetical Means as: Total Variance – EPV = .784 – .66 = **.124**.

Solution 3.1.8: Let m be the mean claim frequency for an insured. Then $h(m) = 1/10$ on $[0, 10]$. The mean severity for a risk is r, since that is the mean for the given exponential distribution. Therefore, for a given insured the mean pure premium is mr. The first moment of the hypothetical mean pure premiums is (since the frequency and severity distributions are independent):

$$\int_{m=0}^{m=10}\int_{r=0}^{r=2} mrg(r)h(m)dr\,dm = \int_{m=0}^{m=10} m/10\,dm \int_{r=0}^{r=2} r(r/2)dr$$

$$= (5)(4/3) = 6.667$$

The second moment of the hypothetical mean pure premiums is (since the frequency and severity distributions are independent):

$$\int_{m=0}^{m=10}\int_{r=0}^{r=2} m^2r^2g(r)h(m)dr\,dm = \int_{m=0}^{m=10} m^2/10\,dm \int_{r=0}^{r=2} r^2(r/2)dr$$

$$= (100/3)(2) = 66.667$$

Therefore, the variance of the hypothetical mean pure premiums is $66.667 - 6.667^2 =$ **22.22**.

Comment: Note that when frequency and severity are independent, the second moment of their product is equal to the product of their second moments. The same is *not* true for variances.

Solution 3.1.9: For a Bernoulli the process variance is $pq = p(1-p)$. For example for Die A_1, the process variance $= (2/6)(1-2/6) = 2/9 = .2222$.

Type of Die	Bernoulli Parameter	A Priori Probability	Process Variance
A1	0.3333	0.50	0.2222
A2	0.5000	0.50	0.2500
Average			**0.2361**

Solution 3.1.10:

Type of Die	A Priori Probability	Mean	Square of Mean
A1	0.50	0.33333	0.11111
A2	0.50	0.50000	0.25000
Average		0.41667	0.18056

Thus, the Variance of the Hypothetical Means= $.18056 - .41667^2$ = **.00695**.

Solution 3.1.11: For spinner B_1 the first moment is $(20)(.6) + (50)(.4) = 32$ and the second moment is $(20^2)(.6) + (50^2)(.4) = 1{,}240$. Thus the process variance is $1{,}240 - 32^2 = 216$. For spinner B_2 the first moment is $(20)(.2) + (50)(.8) = 44$ and the second moment is $(20^2)(.2) + (50^2)(.8) = 2{,}080$. Thus the process variance is $2{,}080 - 44^2 = 144$. Therefore, the expected value of the process variance = $(1/2)(216) + (1/2)(144) =$ **180**.

Type of Spinner	A Priori Probability	Mean	Second Moment	Process Variance
B1	0.50	32	1,240	216
B2	0.50	44	2,080	144
Average				180

Solution 3.1.12:

Type of Spinner	A Priori Probability	Mean	Square of Mean
B1	0.50	32	1,024
B2	0.50	44	1,936
Average		38	1,480

Thus, the Variance of the Hypothetical Means= $1{,}480 - 38^2 = 36$.

Comment: Note that the spinners are chosen independently of the dice, so frequency and severity are independent across risk types. Thus, one can ignore the frequency process in this and the prior question. One can not do so when for example low frequency is associated with low severity.

Solution 3.1.13: For each possible pair of die and spinner use the formula: variance of *p.p.* = $\mu_f \sigma_s^2 + \mu_s^2 \sigma_f^2$.

Die and Spinner	A Priori Chance of Risk	Mean Freq.	Variance of Freq.	Mean Severity	Variance of Sev.	Process Variance of P.P.
A1, B1	0.250	0.333	0.222	32	216	299.6
A1, B2	0.250	0.333	0.222	44	144	478.2
A2, B1	0.250	0.500	0.250	32	216	364.0
A2, B2	0.250	0.500	0.250	44	144	556.0
Mean						**424.4**

Solution 3.1.14:

Die and Spinner	A Priori Chance of Risk	Mean Freq.	Mean Severity	Mean Pure Premium	Square of Mean P.P.
A1, B1	0.250	0.333	32	10.667	113.778
A1, B2	0.250	0.333	44	14.667	215.111
A2, B1	0.250	0.500	32	16.000	256.000
A2, B2	0.250	0.500	44	22.000	484.000
Mean				15.833	267.222

Thus, the Variance of the Hypothetical Means = $267.222 - 15.833^2 =$ **16.53**.

Solution 3.1.15: For the Poisson the process variance is the equal to the mean. The expected value of the process variance is the weighted average of the process variances for the individual types, using the a priori probabilities as the weights. The EPV of the frequency = (40%)(6) + (35%)(7) + (25%)(9) = 7.10.

Type	A Priori Probability	Poisson Parameter	Process Variance
1	40%	6	6
2	35%	7	7
3	25%	9	9
Average			**7.10**

Solution 3.1.16: One computes the first and 2nd moments of the mean frequencies as follows:

Type	A Priori Probability	Poisson Parameter	Mean Frequency	Square of Mean Freq.
1	40%	6	6	36
2	35%	7	7	49
3	25%	9	9	81
Average			7.10	51.80

Then the variance of the hypothetical mean frequencies $= 51.80 - 7.10^2 = \mathbf{1.39}$.

Comment: Using the solution to this question and the previous question, as explained in the next section, the Bühlmann Credibility parameter for frequency is $K = \text{EPV}/\text{VHM} = 7.10/1.39 = 5.11$. The Bühlmann Credibility applied to the observation of the frequency for E *exposures* would be: $Z = E/(E + 5.11)$.

Solution 3.1.17: One has to weight together the process variances of the severities for the individual types using the chance that a claim came from each type. The chance that a claim came from an individual type is proportional to the product of the a priori chance of an insured being of that type and the mean frequency for that type.

As per the Appendix, parameterize the Exponential with mean $1/\lambda$. For type 1, the process variance of the Exponential severity is $1/\lambda^2 = 1/.01^2 = 10{,}000$. (For the Exponential Distribution, the variance is the square of the mean.) Similarly for type 2 the process variance for the severity is $1/.008^2 = 15{,}625$. For type 3 the process variance for the severity is $1/.005^2 = 40{,}000$.

The mean frequencies are: 6, 7, and 9. The a priori chances of each type are: 40%, 35% and 25%. Thus, the weights to use to compute the EPV of the severity are (6)(40%), (7)(35%), (9)(25%) = 2.4, 2.45, 2.25. The expected value of the process variance of the severity is the weighted average of the process variances for the individual types, using these weights. The EPV of the severity = {(2.4)(10,000) + (2.45)(15,625) + (2.25)(40,000)}/(2.4 + 2.45 + 2.25) = 21,448.

A	B	C	D	E	F
Type	A Priori Probability	Mean Frequency	Weights= Col. B × Col. C	Exponential Parameter λ	Process Variance
1	40%	6	2.40	0.01	10,000
2	35%	7	2.45	0.008	15,625
3	25%	9	2.25	0.005	40,000
Average			7.10		**21,448**

Solution 3.1.18: In computing the moments one has to use for each individual type the chance that a claim came from that type. The chance that a claim came from an individual type is proportional to the product of the a priori chance of an insured being of that type and the mean frequency for that type. Thus, the weights to use to compute the moments of the mean severities are: (6)(40%), (7)(35%), (9)(25%) = 2.4, 2.45, 2.25.

A	B	C	D	E	F
Type	A Priori Probability	Mean Frequency	Weights= Col. B × Col. C	Mean Severity	Square of Mean Severity
1	40%	6	2.40	100	10,000
2	35%	7	2.45	125	15,625
3	25%	9	2.25	200	40,000
Average			7.10	140.32	21,448

Then the variance of the hypothetical mean severities = 21,448 – 140.32^2 = **1,758**.

Comment: Using the solution to this question and the previous question, as explained in the next section, the Bühlmann Credibility parameter for severity is $K = \text{EPV}/\text{VHM} = 21{,}448/1{,}758 = 12.20$. The Bühlmann Credibility applied to the observation of the mean severity for N *claims* would be: $Z = N/(N + 12.20)$.

Solution 3.1.19: For type 1 the mean frequency is 6 and the variance of the frequency is also 6. As per the Appendix, parameterize the Exponential with mean $1/\lambda$. For type 1 the mean severity is 100. For type 1, the variance of the Exponential severity is $1/\lambda^2 = 1/.01^2 = 10{,}000$. (For the Exponential Distribution, the variance is the square of the mean.) Thus, since frequency and severity are assumed to be independent, the process variance of the pure premium = (Mean Frequency)(Variance of Severity) + (Mean Severity)^2 (Variance of Frequency) = $(6)(10{,}000) + (100)^2(6) = 120{,}000$. Similarly for type 2 the process variance of the pure premium = $(7)(15{,}625) + (125)^2(7) = 218{,}750$. For type 3 the process variance of the pure premium = $(9)(40{,}000) + (200)^2(9) = 720{,}000$. The expected value of the process variance is the weighted average of the process variances for the individual types, using the a priori probabilities as the weights. The EPV of the pure premium = (40%)(120,000) + (35%)(218,750) + (25%)(720,000) = **304,562**.

Type	A Priori Probability	Mean Frequency	Variance of Frequency	Mean Severity	Variance of Severity	Process Variance
1	40%	6	6	100	10,000	120,000
2	35%	7	7	125	15,625	218,750
3	25%	9	9	200	40,000	720,000
Average						**304,562**

Solution 3.1.20: One has to first compute the mean pure premium for each type. Since frequency and severity are assumed to be independent, the mean pure premium = (Mean Frequency)(Mean

Severity). Then one computes the first and second moments of the mean pure premiums as follows:

Type	A Priori Probability	Mean Frequency	Mean Severity	Mean Pure Premium	Square of Pure Premium
1	40%	6	100	600	360,000
2	35%	7	125	875	765,625
3	25%	9	200	1,800	3,240,000
Average				996.25	1,221,969

Then the variance of the hypothetical mean pure premiums = $1{,}221{,}969 - 996.25^2 =$ **229,455**.

Comment: Using the solution to this question and the previous question, as explained in the next section, the Bühlmann Credibility parameter for the pure premium is $K = \text{EPV}/\text{VHM} = 304{,}562/229{,}455 = 1.33$. The Bühlmann Credibility applied to the observation of the pure premium for E *exposures* would be: $Z = E/(E + 1.33)$.

Solution 3.1.21: The process variance for a binomial distribution is $npq = 5p(1-p)$. $\text{EPV} = \int_0^1 5p(1-p)dp = (5p^2/2 - 5p^3/3)\big|_0^1$ $=$ **5/6**.

Solution 3.1.22: The mean for the binomial is $\mu = np = 5p$. $\text{VHM} = E[\mu^2] - (E[\mu])^2$. The first moment is $E[5p] = \int_0^1 5p\,dp = 5/2$. The second moment is $E[(5p)^2] = \int_0^1 (5p)^2 dp = 25/3$. So, $\text{VHM} = E[(5p)^2] - (E[5p])^2 = 25/3 - 25/4 =$ **25/12**.

Solution 3.2.1: The Bühlmann Credibility parameter is: $K =$ (The Expected Value of the Process Variance)/(The Variance of the Hypothetical Means) $= 100/8 = 12.5$. $Z = N/(N+K) = 20/(20+12.5) =$ **61.5%**.

Solution 3.2.2: $Z = N/(N+K)$, therefore, $K = N(1/Z) - N = 5(1/.7) - 5 =$ **2.14**.

Solution 3.2.3:

Type of Die	A Priori Chance of This Type of Die	Process Variance	Mean Die Roll	Square of Mean Die Roll
4-sided	0.333	1.250	2.5	6.25
6-sided	0.333	2.917	3.5	12.25
8-sided	0.333	5.250	4.5	20.25
Overall		3.1389	3.50	12.91667

The variance of the hypothetical means $= 12.91667 - 3.5^2 = .6667$. K = EPV/VHM = 3.1389/.6667 = 4.71. $Z = (1/1 + 4.71) = .175$. The a priori estimate is 3.5 and the observation is 5, so the new estimate is (.175)(5) + (.825)(3.5) = **3.76**.

Solution 3.2.4: Expected Value of the Process Variance = .0833.

Variance of the Hypothetical Means $= .4167 - .5^2 = .1667$.

Type of Urn	A Priori Probability	Mean for This Type Urn	Square of Mean of This Type Urn	Process Variance
1	0.3333	0	0	0
2	0.3333	1	1	0.00000
3	0.3333	0.5	0.25	0.25000
Average		0.5	0.4167	0.0833

K = EPV/VHM = .0833/.1667 = .5 Thus for $N = 5$, $Z = 5/(5 + .5) = 90.9\%$. The observed mean is 0 and the a priori mean is .5, therefore, the new estimate = (0)(.909) + (.5)(1 − .909) = **.0455**.

Solution 3.2.5: As computed in the solution to the previous question, for 5 observations $Z = 90.9\%$ and the a priori mean is

.5. Since the observed mean is $2/5 = .4$, the new estimate is: $(.4)(.909) + (.5)(1 - .909) =$ **.4091**.

Solution 3.2.6: For example, the second moment of Urn II is $(.7)(1{,}000^2) + (.3)(2{,}000^2) = 1{,}900{,}000$. The process variance of Urn II $= 1{,}900{,}000 - 1{,}300^2 = 210{,}000$.

Type of Urn	A Priori Probability	Mean	Square of Mean	Second Moment	Process Variance
I	0.8000	1,100	1,210,000	1,300,000	90,000
II	0.2000	1,300	1,690,000	1,900,000	210,000
Average		1,140	1,306,000		114,000

Thus, the expected value of the process variance = 114,000, and the variance of the hypothetical means is: $1{,}306{,}000 - 1{,}140^2 = 6{,}400$.

Thus, the Bühlmann Credibility parameter is $K = \text{EPV}/\text{VHM} = 114{,}000/6{,}400 = 17.8$. Thus for 5 observations $Z = 5/(5 + 17.8) = 21.9\%$. The prior mean is \$1,140 and the observation is $8{,}000/5 = \$1{,}600$.

Thus the new estimate is: $(.219)(1{,}600) + (1 - .219)(1{,}140) =$ **\$1,241**.

Solution 3.2.7: For Risk A the mean is $(.07)(100) + (.03)(500) = 22$ and the second moment is $(.07)(100^2) + (.03)(500^2) = 8{,}200$. Thus, the process variance for Risk A is $8{,}200 - 22^2 = 7{,}716$. Similarly for Risk B the mean is 130 and the second moment is 53,000. Thus, the process variance for Risk B is $53{,}000 - 130^2 = 36{,}100$. For Risk C the mean is 218 and the second moment is 95,800. Thus, the process variance for Risk C is $95{,}800 - 218^2 = 48{,}276$. Thus, the expected value of the process variance $= (1/3)(7{,}716) + (1/3)(36{,}100) + (1/3)(48{,}276) = 30{,}697$.

Risk	A Priori Chance of Risk	Mean	Square of Mean
A	0.333	22	484
B	0.333	130	16,900
C	0.333	218	47,524
Mean		123.33	21,636

Thus, the Variance of the Hypothetical Means = $21{,}636 - 123.33^2 = 6{,}426$.

Therefore, the Bühlmann Credibility Parameter for pure premium = K = EPV/VHM = 30,697/6,426 = 4.78. Thus, the credibility for 1 observation is $1/(1+K) = 1/5.78 = .173$. The a priori mean is 123.33. The observation is 500. Thus, the estimated aggregate losses are: (.173)(500) + (1 − .173)(123.33) = **188.5**.

Solution 3.2.8: For Die A the mean is $(1+1+1+2+3+4)/6 = 2$ and the second moment is $(1+1+1+4+9+16)/6 = 5.3333$. Thus the process variance for Die A is $5.3333 - 2^2 = 1.3333$. Similarly for Die B the mean is 2.3333 and the second moment is 6.3333. Thus, the process variance for Die B is $6.333 - 2.333^2 = .889$. The mean of Die C is 2.6667. The process variance for Die C is: $\{(1-2.6667)^2 + (2-2.6667)^2 + (3)(3-2.6667)^2 + (4-2.6667)^2\}/6 = .889$, the same as Die B. The mean of Die D is 3. The process variance for Die D is: $\{(1-3)^2 + (2-3)^2 + (3-3)^2 + (3)(4-3)^2\}/6 = 1.333$, the same as Die A. Thus, the expected value of the process variance = (1/4)(1.333) + (1/4)(.889) + (1/4)(.889) + (1/4)(1.333) = 1.111.

Die	A Priori Chance of Die	Mean	Square of Mean
A	0.250	2.0000	4.0000
B	0.250	2.3333	5.4443
C	0.250	2.6667	7.1113
D	0.250	3.0000	9.0000
Mean		2.5000	6.3889

Thus, the Variance of the Hypothetical Means = $6.3889 - 2.5^2 =$ .1389.

Therefore, the Bühlmann Credibility Parameter = K = EPV/VHM = 1.111/.1389 = 8.0. Thus, the creibility for 5 observations is $5/(5+K) = 5/13$. The a priori mean is 2.5. The observed mean is $(2+3+1+2+4)/5 = 2.4$. Thus, the estimated future die roll is: $(5/13)(2.4) + (1 - 5/13)(2.5) =$ **2.462**.

Solution 3.2.9: For each Poisson, the process variance is the mean. Therefore, Expected Value of the process variance = (.6)(.05) + (.3)(.1) + (.1)(.2) = .08 = Overall mean frequency. The expected value of the square of the mean frequencies is .0085. Therefore, the Variance of the hypothetical mean frequencies = $.0085 - .08^2 = .0021$. Alternately, Variance of the hypothetical mean frequencies = $(.6)(.03^2) + (.3)(.02^2) + (.1)(.12^2) = .0021$. Therefore, K = EPV/VHM = .08/.0021 = 38.1. $Z = 5/(5 + 38.1) = 11.6\%$. Estimated frequency = (11.6%)(.2)+ (88.4%)(.08) = **.0939**.

Type of Driver	A Priori Chance of This Type of Driver	Mean Annual Claim Frequency	Square of Mean Claim Frequency	Poisson Process Variance
Good	0.6	0.05	0.0025	0.05
Bad	0.3	0.1	0.0100	0.1
Ugly	0.1	0.2	0.0400	0.2
Average		0.080	0.0085	0.080

Solution 3.2.10: One needs to figure out for each type of driver a single observation of the risk process, in other words for the observation of a single claim, the process variance of the average size of a claim. Process variances for the Pareto Distributions are $\lambda^2\alpha/\{(\alpha-1)^2(\alpha-2)\}$, so the process variances are: 10.42, 22.22, and 75 million. The probability weights are the product of claim frequency and the a priori frequency of each type of driver: (.6)(.05), (.3)(.10), (.1)(.20). The probabilities that a claim came

from each of the types of drivers are the probability weights divided by the their sum: .375, .375, .25. Thus, the weighted average process variance of the severity is: (10.42 million)(.375) + (22.22 million)(.375) + (75 million)(.25) = **30.98 million**.

Type of Driver	A Priori Chance of This Type of Driver	Average Claim Frequency	Probability Weight For Claim	Probability For Claim	Alpha	Process Variance Claim Severity
Good	0.6	0.05	0.030	0.375	5	1.042×10^7
Bad	0.3	0.1	0.030	0.375	4	2.222×10^7
Ugly	0.1	0.2	0.020	0.25	3	7.500×10^7
Average			0.080	1.000		$\mathbf{3.098 \times 10^7}$

Comment: A claim is more likely to be from a Good Driver since there are many Good Drivers. On the other hand, a claim is more likely to be from an Ugly Driver, because each such driver produces more claims. Thus, the probability that a claim came from each type of driver is proportional to the product of claim frequency and the a priori frequency of each type of driver. The (process) variances for the Pareto Distribution follow from the moments given in the Appendix.

Solution 3.2.11: Average severities for the Pareto Distributions are: $\lambda/(\alpha - 1)$ = 2,500, 3,333 and 5,000. Probability weights are: (.60)(.05), (.30)(.10), (.10)(.20). The overall average severity is 3437.5. Average of the severity squared is: (.375)(6.25 million) + (.375)(11.11 million) + (.25)(25 million) = 12.76 million. Therefore, the variance of the hypothetical mean severities = (12.76 million) – (3437.5^2) = **.94 million**.

Type of Driver	A Priori Chance of This Type of Driver	Average Claim Frequency	Probability Weight For Claim	Probability For Claim	Alpha	Average Claim Severity	Square of Average Claim Severity
Good	0.6	0.05	0.030	0.375	5	2,500	6.250×10^6
Bad	0.3	0.1	0.030	0.375	4	3,333	1.111×10^7
Ugly	0.1	0.2	0.020	0.250	3	5,000	2.500×10^7
Average			0.080	1.000		3,437.5	1.276×10^7

Solution 3.2.12: Using the solutions to the previous two questions, $K = \text{EPV}/\text{VHM} = 31/.94 = 32.8$. $Z = 1/(1+32.8) = 1/33.8$. New estimate $= (1/33.8)25{,}000 + [1-(1/33.8)]3{,}437.5 =$ **\$4,075**.

Solution 3.2.13: For each type of driver one uses the formula: variance of $p.p. = \mu_f\sigma_s^2 + \mu_s^2\sigma_f^2$. In the case of a Poisson frequency $\mu_f = \sigma_f^2$ and: variance of $p.p.$ = (mean frequency)(the second moment of the severities).

For the Pareto, the second moment $= 2\lambda^2/\{(\alpha-1)(\alpha-2)\}$.

A	B	C	D	E	F
Type of Driver	A Priori Chance of This Type of Driver	Claim Frequency	Alpha	Expected Value of Square of Claim Sizes	Variance of P.P. Product of Columns C & E
Good	0.6	0.05	5	$1.667e+7$	8.333×10^5
Bad	0.3	0.1	4	$3.333e+7$	3.333×10^6
Ugly	0.1	0.2	3	$1.000e+8$	2.000×10^7
Average					$\mathbf{3.500 \times 10^6}$

Solution 3.2.14: The overall average pure premium = (.6)(125) + (.3)(333.3) + (.1)(1,000) = 275. The average of the squares of the hypothetical mean pure premiums is: $(.6)(125^2) + (.3)(333.3^2) + (.1)(1{,}000^2) = 142{,}708$. Therefore, the variance of the hypothetical pure premiums $= 142{,}708 - (275^2) =$ **67,083**.

Type of Driver	A Priori Chance of This Type of Driver	Average Claim Frequency	Alpha	Average Claim Severity	Average Pure Premium	Square of Average Pure Premium
Good	0.6	0.05	5	2,500	125.0	15,625
Bad	0.3	0.1	4	3,333	333.3	111,111
Ugly	0.1	0.2	3	5,000	1,000.0	1,000,000
Average					275.0	142,708

Solution 3.2.15: The observed pure premium is \$25,000/5 = \$5,000. Using the results of the previous two questions, K = 3,500,000/67,083 = 52. $Z = 5/(5+52) = 8.8\%$. Estimated pure premium = (8.8%)(\$5,000) + (1 – 8.8%)(\$275) = **\$691**.

Solution 3.2.16:

Class	A Priori Probability	Mean for This Class	Square of Mean of This Class	Process Variance
1	0.4000	0.2	0.04	0.16
2	0.6000	0.3	0.09	0.21
Average		0.26	0.0700	0.1900

Expected Value of the Process Variance = .19.

Variance of the Hypothetical Means = $.070 - .26^2 = .0024$.

K = EPV/VHM = .19/.0024 = 79.2 Thus for $N = 5$, $Z = 5/(5+79.2) =$ **5.94%**.

Solution 3.2.17: The hypothetical mean pure premiums are (.1667)(4) and (.8333)(2); which are 2/3 and 5/3. Since the two classes have the same number of risks the overall mean is 7/6 and the variance of the hypothetical mean pure premiums between classes is: $[(2/3 - 7/6)^2 + (5/3 - 7/6)^2]/2 = 1/4$.

Each class is homogeneous and the stated data are the process variance for a risk from each class. For each type of risk, the process variance of the pure premiums is given by: $\mu_f \sigma_s^2 + \mu_f^2 \sigma_s^2$. For Class A, that is: $(.1667)(20) + (4^2)(.1389) = 5.5564$. For Class B, that is: $(.8333)(5) + (2^2)(.1389) = 4.7221$. Since the classes have the same number of risks, the Expected Value of the Process Variance = (.5)(5.5564) + (.5)(4.7221) = 5.139. Thus K = EPV/VHM = 5.139/.25 = 20.56. $Z = N/(N+K) = 4/(4+20.56) =$ **.163**.

Solution 3.2.18: The prior estimate is the overall mean of $7/6$. The observation is .25. Thus, the new estimate is $(.163)(.25) + (7/6)(1 - .163) =$ **1.017**.

Comment: Uses the solution of the previous question.

Solution 3.2.19: Expected Value of the Process Variance = $E[v] = 8$.

Variance of the Hypothetical Means = $\text{Var}[m] = 4$.

$K = \text{EPV}/\text{VHM} = 8/4 = 2$. So, $Z = 3/(3 + K) = 3/(3 + 2) = 3/5 =$ **.6**.

Solution 3.3.1: 1. True. 2. False. 3. True.

Solution 3.3.2: The expected value of the process variance is 86.333. The variance of the hypothetical means is $466.67 - 20^2 = 66.67$.

Type of Marksman	A Priori Chance of This Type of Marksman	Mean	Square of Mean	Standard Deviation	Process Variance
A	0.333	10	100	3	9
B	0.333	20	400	5	25
C	0.333	30	900	15	225
Average		20.000	466.667		86.333

$K = \text{EPV}/\text{VHM} = 86.333/66.67 = 1.295$.

$Z = N/(N + K) = 2/(2 + 1.295) = .607$. The average observation is $(10 + 14)/2 = 12$. The a priori mean = 20. Thus, the new estimate = $(.607)(12) + (1 - .607)(20) =$ **15.14**.

Solution 3.3.3: The expected value of the process variance is 13. The variance of the hypothetical means is $1 - 0^2 = 1$. Therefore, $K = \text{EPV}/\text{VHM} = 13/1 = 13$. $Z = 1/(1 + 13) = 1/14$. New Estimate = $(4)(1/14) + (0)(1 - 1/14) = 2/7 =$ **.286**.

Shooter	A Priori Probability	Process Variance	Mean	Square of Mean
A	0.5	1	1	1
B	0.5	25	−1	1
Average		13	0	1

Solution 3.3.4: The EPV is $(1/2)(1+25) = 13$. The VHM is $(1/2)[(1-0)^2 + (-1-0)^2] = 1$. Therefore, $K = \text{EPV}/\text{VHM} = 13/1 = 13$. Since there are three observations, $n = 3$ and $Z = n/(n+K) = 3/(3+13) = 3/16$. The average position of the three shots is $(1/3)(2+0+1) = 1$. So, Estimate $= (1)(3/16) + (0)(1-3/16) =$ **3/16** = **.188**.

Solution 4.1.1: The probability of picking a colorblind person out of this population is (5%)(10%) + (1/4%)(90%) = .725%. The chance of a person being both colorblind and male is: (5%)(10%) = .5%. Thus the (conditional) probability that the colorblind person is a man is: .5%/.725% = **69.0%**.

Solution 4.1.2: Taking a weighted average, the a priori chance of a black ball is 8.8%.

A	B	C	D
Type	A Priori Probability	% Black Balls	Col. B × Col. C
I	0.4	0.05	0.020
II	0.3	0.08	0.024
III	0.2	0.13	0.026
IV	0.1	0.18	0.018
Sum	1		**0.088**

Solution 4.1.3: $P[\text{Urn} = \text{I} \mid \text{Ball} = \text{Black}] = P[\text{Urn} = \text{I and Ball} = \text{Black}]/P[\text{Ball} = \text{Black}] = .020/.088 =$ **22.7%**.

Solution 4.1.4: P[Urn = II | Ball = Black] = P[Urn = II and Ball = Black]/P[Ball = Black] = .024/.088 = **27.3%**.

Solution 4.1.5: P[Urn = III | Ball = Black] = P[Urn = III and Ball = Black]/P[Ball = Black] = .026/.088 = **29.5%**.

Solution 4.1.6: P[Urn = IV | Ball = Black] = P[Urn = IV and Ball = Black]/P[Ball = Black] = .018/.088 = **20.5%**.

Comment: The conditional probabilities of the four types of urns add to unity.

Solution 4.1.7: Using the solutions to the previous problems, one takes a weighted average using the posterior probabilities of each type of urn: (22.7%)(.05) + (27.3%)(.08) + (29.5%)(.13) + (20.5%)(.18) = **.108**.

Comment: This whole set of problems can be usefully organized into a spreadsheet:

A	B	C	D	E	F
Type	A Priori Probability	% Black Balls	Probability Weights = Col. B × Col. C	Posterior Probability	Col. C × Col. E
I	0.4	0.05	0.0200	0.227	0.011
II	0.3	0.08	0.0240	0.273	0.022
III	0.2	0.13	0.0260	0.295	0.038
IV	0.1	0.18	0.0180	0.205	0.037
Sum			0.0880	1.000	**0.108**

This is a simple example of Bayesian Analysis, which is covered in the next section.

Solution 4.1.8: There are the following 10 equally likely possibilities such that $Y \geq 9$: (3,6), (4,5), (4,6), (5,4), (5,5), (5,6), (6,3), (6,4), (6,5), (6,6). Of these, 3 have $X = 5$, so that Prob[$X = 5 \mid Y \geq 9$] = Prob[$X = 5$ and $Y \geq 9$]/Prob[$Y \geq 9$] = (3/36)/(10/36) = 3/10 = **.3**.

Solution 4.1.9: There are the following 10 equally likely possibilities such that $X + V \geq 9$: (3,6), (4,5), (4,6), (5,4), (5,5), (5,6), (6,3), (6,4), (6,5), (6,6). Of these one has $X = 3$, two have $X = 4$, three have $X = 5$, and four have $X = 6$. Therefore, $E[X \mid Y \geq 9] = \{(1)(3) + (2)(4) + (3)(5) + (4)(6)\}/10 =$ **5.0**. For those who like diagrams:

X	1	2	3	4	5	6	Possibilities	Conditional Density Function of X given that $X + V \geq 9$
1							0	
2							0	
3						x	1	1/10
4					x	x	2	2/10
5				x	x	x	3	3/10
6			x	x	x	x	4	4/10

$E[X \mid Y \geq 9] = \sum iP[X = i \mid Y \geq 9] = (1)(0) + (2)(0) + (3)(.1) + (4)(.2) + (5)(.3) + (6)(.4) = 5.0.$

Solution 4.1.10: The chance that a driver accident-free is: (40%)(80%)+(25%)(85%)+(20%)(90%)+(15%)(95%) = 85.5%. The chance that a driver is both accident-free and from Boston is (40%)(80%) = 32%. Thus the chance this driver is from Boston is 32%/85.5% = **37.4%**.

Comment: Some may find to helpful to assume for example a total of 100,000 drivers.

Solution 4.1.11: The chance that a driver has had an accident is: (40%)(20%) + (25%)(15%) + (20%)(10%) + (15%)(5%) = 14.5%. The chance that a driver both has had an accident and is from Pittsfield is (15%)(5%) = .75%. Thus, the chance this driver is from Pittsfield is: .75%/14.5% = **.052**.

Comment: Note that the chances for each of the other cities are: {(40%)(20%), (25%)(15%), (20%)(10%)}/14.5%. You should confirm that the conditional probabilities for the four cities sum to 100%.

Solution 4.1.12: If Stu knows the answer, then the chance of observing a correct answer is 100%. If Stu doesn't know the answer to a question then the chance of observing a correct answer is 20%.

A	B	C	D	E
Type of Question	A Priori Chance of This Type of Question	Chance of the Observation	Prob. Weight = Product of Columns B & C	Posterior Chance of This Type of Question
Stu Knows	0.620	1.0000	0.6200	**89.08%**
Stu Doesn't Know	0.380	0.2000	0.0760	10.92%
Overall			0.696	1.000

Solution 4.2.1: (20%)(60%)+(30%)(25%)+(40%)(15%) = **25.5%**.

Comment: Since there are only two possible outcomes, the chance of observing no claim is: 1 – .255 = .745.

Solution 4.2.2: P(Type A | no claim) = P(no claim | Type A) P(Type A)/P(no claim) = (.8)(.6)/.745 = **64.43%**.

Solution 4.2.3: (.7)(.25)/.745 = **23.49%**.

Solution 4.2.4: (.6)(.15)/.745 = **12.08%**.

Solution 4.2.5:

A	B	C	D	E	F
Type of Risk	A Priori Chance of This Type of Risk	Chance of the Observation	Prob. Weight = Product of Columns B & C	Posterior Chance of This Type of Risk	Mean Annual Freq.
A	0.6	0.8	0.480	64.43%	0.20
B	0.25	0.7	0.175	23.49%	0.30
C	0.15	0.6	0.090	12.08%	0.40
Overall			0.745	1.000	**24.77%**

Solution 4.2.6:

A	B	C	D	E	F
Type of Risk	A Priori Chance of This Type of Risk	Chance of the Observation	Prob. Weight = Product of Columns B & C	Posterior Chance of This Type of Risk	Mean Annual Freq.
A	0.6	0.2	0.120	47.06%	0.20
B	0.25	0.3	0.075	29.41%	0.30
C	0.15	0.4	0.060	23.53%	0.40
Overall			0.255	1.000	**27.65%**

Solution 4.2.7:

A	B	C	D	E	F
Type of Risk	A Priori Chance of This Type of Risk	Chance of the Observation	Prob. Weight = Product of Columns B & C	Posterior Chance of This Type of Risk	Mean Annual Freq.
A	0.6	0.2048	0.123	48.78%	0.20
B	0.25	0.3087	0.077	30.64%	0.30
C	0.15	0.3456	0.052	20.58%	0.40
Overall			0.252	1.000	**27.18%**

For example, if one has a risk of Type B, the chance of observing 2 claims in 5 years is given by (a Binomial Distribution): $(10)(.3^2)(.7^3) = .3087$.

Solution 4.2.8: Bayesian Estimates are in balance; the sum of the product of the a priori chance of each outcome times its posterior Bayesian estimate is equal to the a priori mean. The a priori mean is $(5/8)(1) + (2/8)(4) + (1/8)(16) = 3.625$. Let $E[X_2 \mid X_1 = 16] = y$. Then setting the sum of the chance of each outcome times its posterior mean equal to the a priori mean: $(5/8)(1.4) + (2/8)(3.6) + (1/8)(y) = 3.625$. Therefore, $y = \mathbf{14.8}$.

Solution 4.2.9:

A	B	C	D	E	F
Type of Urn	A Priori Chance of This Type of Urn	Chance of the Observation	Prob. Weight = Product of Columns B & C	Posterior Chance of This Type of Urn	Mean Draw from Urn
I	0.8000	0.1000	0.0800	0.5714	1100
II	0.2000	0.3000	0.0600	0.4286	1300
Overall			0.140	1.000	**1186**

Solution 4.2.10:

A	B	C	D	E	F
Type of Urn	A Priori Chance of This Type of Urn	Chance of the Observation	Prob. Weight = Product of Columns B & C	Posterior Chance of This Type of Urn	Mean Draw from Urn
I	0.8000	0.2430	0.1944	0.6879	1100
II	0.2000	0.4410	0.0882	0.3121	1300
Overall			0.283	1.000	**1162**

For example, the chance of picking 2 @ \$1,000 and 1 @ \$2,000 from Urn II is given by $f(2)$ for a Binomial Distribution with $n = 3$ and $p = .7$: $(3)(.7^2)(.3) = .4410$.

Solution 4.2.11: The chance of observing 3 claims for a Poisson is $e^{-\theta}\theta^3/3!$ Therefore, for example, the chance of observing 3 claims for a risk of type 1 is: $e^{-.4}(.4^3)/6 = .00715$.

A	B	C	D	E	F
Type	A Priori Probability	Chance of the Observation	Prob. Weight = Product of Columns B & C	Posterior Chance of This Type of Risk	Mean Annual Freq.
1	70%	0.00715	0.005005	39.13%	0.4
2	20%	0.01976	0.003951	30.89%	0.6
3	10%	0.03834	0.003834	29.98%	0.8
Overall			0.012791	1.000	**0.5817**

Solution 4.2.12: The density for a Normal Distribution with mean μ and standard deviation σ is given by $f(x) = \exp(-.5\{(x-\mu)/\sigma\}^2)/\{\sigma(2\pi)^{.5}\}$. Thus, the density function at 14 for Marksman A is $\exp(-.5\{(14-10)/3\}^2)/\{3(2\pi)^{.5}\} = .0547$.

A	B	C	D	E	F	G	H	I
Marks-man	Mean	Standard Deviation	A Priori Chance of This Type of Marksman	Chance of the Observing 10	Chance of the Observing 14	Chance of the Observation	Prob. Weight = Product of Columns D & G	Posterior Chance of This Type of Marksman
A	10	3	0.333	0.1330	0.0547	0.007270	0.002423	92.56%
B	20	5	0.333	0.0108	0.0388	0.000419	0.000140	**5.34%**
C	30	15	0.333	0.0109	0.0151	0.000165	0.000055	2.10%
Overall							0.002618	**1.000**

Reading from the table above, we see that the chance that it was marksman B is **5.34%**.

Solution 4.2.13: Use the results of the previous question to weight together the prior means:

Marksman	Posterior Chance of This Type of Risk	A Priori Mean
A	0.9256	10
B	0.0534	20
C	0.0210	30
Overall		**10.954**

Solution 4.3.1: Use equation [4.3.3]: $f_\theta(\theta \mid X = x) = f_X(x \mid \theta) f_\theta(\theta)/f_X(x)$ where x is the event of a claim in each of two years. The probability of one claim in one year is θ, so the probability of a claim in each of two years given θ is $f_X(x \mid \theta) = \theta^2$. By definition, $f_\theta(\theta) = 1$ for $0 \le \theta \le 1$. The last piece is $f_X(x) = \int_0^1 f_X(x \mid \theta) f(\theta) d\theta = \int_0^1 \theta^2 \cdot 1\, d\theta = \theta^3/3 \,|_0^1 = 1/3$. The answer is: $f_\theta(\theta \mid X = x) = \theta^2 \cdot 1/(1/3) = \mathbf{3\theta^2}$ **for** $\mathbf{0 \le \theta \le 1}$.

Solution 4.3.2: Use equation [4.3.3]: $f_\theta(\theta \mid X = x) = f_X(x \mid \theta) f_\theta(\theta)/f_X(x)$ where x the event of a claim in each of three years. The probability of one claim in one year is θ, so the probability of a claim in each of three years given θ is $f_X(x \mid \theta) = \theta^3$. By definition, $f_\theta(\theta) = 1$ for $0 \le \theta \le 1$. The last piece is $f_X(x) = \int_0^1 f_X(x \mid \theta) f(\theta) d\theta = \int_0^1 \theta^3 \cdot 1\, d\theta = \theta^4/4 \,|_0^1 = 1/4$. The answer is: $f_\theta(\theta \mid X = x) = \theta^3 \cdot 1/(1/4) = \mathbf{4\theta^3}$ **for** $\mathbf{0 \le \theta \le 1}$.

Solution 4.3.3: Assuming a given value of p, the chance of observing one success in three trials is $3p(1-p)^2$. The prior distribution of p is: $g(p) = 1, 0 \le p \le 1$. By Bayes Theorem, the posterior distribution of p is proportional to the product of the chance of the observation and the prior distribution: $3p(1-p)^2$. Thus the posterior distribution of p is proportional to $p - 2p^2 + p^3$. (You can keep the factor of 3 and get the same result working instead with $3p - 6p^2 + 3p^3$.) The integral of $p - 2p^2 + p^3$ from 0 to 1 is $1/2 - 2/3 + 1/4 = 1/12$. Thus the posterior distribution of p is $12(p - 2p^2 + p^3)$. (The integral of the posterior distribution has to be unity. In this case dividing by $1/12$; i.e., multiplying by 12, will make it so.) The posterior chance of p in [.3,.4] is:

$$12\int_{p=.3}^{.4}(p - 2p^2 + p^3)dp = 6p^2 - 8p^3 + 3p^4 \Big]_{p=.3}^{.4}$$

$$= .42 - .296 + .0525 = \mathbf{.1765}.$$

Solution 4.3.4: From the solution to the prior question, the posterior distribution of p is: $12(p - 2p^2 + p^3)$. The mean of this posterior distribution is:

$$12\int_{p=0}^{1}(p-2p^2+p^3)p\,dp = 4p^3-6p^4+(12/5)p^5\Big]_{p=0}^{1} = .4.$$

The chance of a success on the fourth trial is $E[p] =$ **.4**.

Solution 4.3.5: Given x, the chance of observing three successes is x^3. The a priori distribution of x is $f(x)=1,\ 0\le x\le 1$. By Bayes Theorem, the posterior density is proportional to the product of the chance of the observation and the a priori density function. Thus the posterior density is proportional to x^3 for $0\le x\le 1$. Since the integral from zero to one of x^3 is $1/4$, the posterior density is $4x^3$. (The posterior density has to integrate to unity.) Thus the posterior chance that $x<.9$ is the integral of the posterior density from 0 to .9, which is $.9^4 =$ **.656**. Alternately, by Bayes Theorem (or directly from the definition of a conditional distribution): $\Pr[x<.9\,|\,3 \text{ successes}] = \Pr[3 \text{ successes}\,|\,x<.9]\Pr[x<.9]/\Pr[3 \text{ successes}] = \Pr[3 \text{ successes and } x<.9]/\Pr[3 \text{ successes}] =$

$$\int_{x=0}^{.9}x^3f(x)dx\Big/\int_{x=0}^{1}x^3f(x)dx = \int_{x=0}^{.9}x^3dx\Big/\int_{x=0}^{1}x^3dx$$
$$= \{(.9^4)/4\}/\{(1^4)/4\} = .656.$$

Solution 4.3.6:

$$\int_0^\infty e^{-\theta}f(\theta)d\theta = \int_0^\infty e^{-\theta}36\theta e^{-6\theta}d\theta = 36\int_0^\infty \theta e^{-7\theta}d\theta$$
$$= (36)(\Gamma(2)/7^2) = (36)(1/49) = \mathbf{.735}.$$

Comment: Note that $\int_0^\infty t^{\alpha-1}e^{-\lambda t}dt = \Gamma(\alpha)\lambda^{-\alpha}$. This follows from the fact that the Gamma Distribution as per the Appendix is in fact a distribution function, so that its density integrates to unity. In this case $\alpha=2$ and $\lambda=7$.

Solution 4.3.7:

$$P(Y=0)=\int P(Y=0\,|\,\theta)f(\theta)d\theta=\int e^{-\theta}f(\theta)d\theta.$$

For the first case, $f(\theta)=1/2$ for $0\le\theta\le 2$

$$P(Y=0)=\int_{\theta=0}^{2}e^{-\theta}/2\,d\theta=(1-e^{-2})/2=.432.$$

For the second case, $f(\theta)=e^{-\theta}$ for $\theta>0$ and

$$P(Y=0)=\int_{\theta=0}^{\infty}e^{-2\theta}d\theta=1/2.$$

For the third case, $P(Y=0)=e^{-1}=.368$.

In the first and third cases $P(Y=0)<.45$.

Comment: Three separate problems in which you need to calculate $P(Y=0)$ given $f(\theta)$ and three different conditional distributions $P(Y=y\,|\,\theta)$.

Solution 5.1.1: $\Gamma(5+1)/8^{5+1}=5!/8^6=\mathbf{4.58\times 10^{-4}}$.

Solution 5.1.2: $\lambda^{\alpha}x^{a-1}e^{-\lambda x}/\Gamma(\alpha)=(.1^3)8^2e^{-.8}/\Gamma(3)=\mathbf{.0144}$.

Solution 5.1.3: Variance $=E[x^2]-(E[x])^2=(\alpha+1)(\alpha)/\lambda^2-(\alpha/\lambda)^2=\alpha/\lambda^2$.

Solution 5.1.4: Mean $=\alpha/\lambda=3.0/1.5=2.0$. Variance $=\alpha/\lambda^2=3.0/(1.5)^2=\mathbf{1.33}$.

Solution 5.2.1: Gamma-Poisson has a Negative Binomial mixed frequency distribution. The Negative Binomial has parameters $k=\alpha=4$ and $p=\lambda/(1+\lambda)=.9$. Thus the chance of 5 claims is

$$\binom{5+k-1}{5}p^k(1-p)^5=\binom{8}{5}.9^4(.1)^5=(56)(.6561)(.00001)$$
$$=\mathbf{.000367}.$$

Solution 5.2.2: The prior distribution is Gamma with $\alpha=2.0$ and $\lambda=3.0$. The number of claims for an insured selected at random is a Negative Binomial with parameters $k=\alpha=2.0$ and $p=$

$\lambda/(\lambda+1) = 3/4$. The mean number is $k(1-p)/p = \alpha/\lambda = \mathbf{2/3}$ and the variance is $k(1-p)/p^2 = \alpha(\lambda+1)/\lambda^2 = 2(3+1)/3^2 = \mathbf{8/9}$.

Solution 5.2.3:

$$k = 2, \qquad p = 0.6 \qquad \text{and} \qquad f(n) = \binom{n+2-1}{n}(.6)^2.4^n.$$

So,

$$\begin{aligned}
P[X > 2] &= 1 - P[X=0] - P[X=1] - P[X=2] \\
&= 1 - f(0) - f(1) - f(2) = 1 - \binom{1}{0}(.6)^2(.4)^0 - \binom{2}{1}.6^2.4^1 \\
&\quad - \binom{3}{2}.6^2.4^2 = 1 - .36 - .288 - .1728 = \mathbf{.1792}.
\end{aligned}$$

Solution 5.2.4: The chance of no claims for a Poisson is $e^{-\lambda}$. We average over the possible values of λ:

$$\begin{aligned}
(1/2)\int_1^3 e^{-\lambda} d\lambda &= (1/2)(-e^{-\lambda})\Big|_1^3 = (1/2)(e^{-1} - e^{-3}) \\
&= (1/2)(.368 - .050) = \mathbf{.159}.
\end{aligned}$$

Solution 5.2.5: (1) Binomial: $\mu = np$ and $\sigma^2 = npq$. Then $\mu = np \geq npq = \sigma^2$ for $q \leq 1$. (2) Negative Binomial: $\mu = k(1-p)/p$ and $\sigma^2 = k(1-p)/p^2$. Then, $\mu = k(1-p)/p \leq k(1-p)/p^2 = \sigma^2$ for $p \leq 1$. Rank for equal means: σ^2 binomial $\leq \sigma^2$ poisson $\leq \sigma^2$ negative binomial.

Solution 5.3.1: The Prior Gamma has parameters $\alpha = 4$ and $\lambda = 2$. The Posterior Gamma has parameters $\alpha' = 4 + 5 = 9$ and $\lambda' = 2 + 3 = 5$.

Solution 5.3.2: For the Gamma-Poisson, if the prior Gamma has parameters $\alpha = 3$, $\lambda = 4$, then the Posterior Gamma has parameters $\alpha = 3 + 1$ and $\lambda = 4 + 2$. Posterior Gamma $= \lambda^{\alpha}\theta^{\alpha-1}e^{-\lambda\theta}/\Gamma(\alpha) = 6^4\theta^3 e^{-6\theta}/(3!) = 216\theta^3 e^{-6\theta}$.

Solution 5.3.3: The prior Gamma has $\alpha = 1$ (an Exponential Distribution) and $\lambda = 1$. The posterior Gamma has $\alpha' =$ prior $\alpha +$ number of claims $= 1 + 0 = 1$ (an Exponential Distribution) and $\lambda' =$ prior $\lambda +$ number of exposures $= 1 + 1 = 2$. That is, the posterior density function is: $2e^{-2\theta}$.

Comment: Given θ, the chance of observing zero claims is $\theta^0 e^{-\theta}/0! = e^{-\theta}$. The posterior distribution is proportional to product of the chance of observation and the a priori distribution of θ: $(e^{-\theta})(e^{-\theta}) = e^{-2\theta}$. Dividing by the integral of $e^{-2\theta}$ from 0 to ∞ gives the posterior distribution: $e^{-2\theta}/(1/2) = 2e^{-2\theta}$.

Solution 5.3.4: For the Gamma-Poisson, the posterior Gamma has shape parameter $\alpha' =$ prior $\alpha +$ number of claims observed $= 50 + 65 + 112 = 227$. For the Gamma Distribution, the mean is α/λ, while the variance is α/λ^2. Thus the coefficient of variation is: $(\text{variance}^{.5})/\text{mean} = \{\alpha/\lambda^2\}^{.5}/\{\alpha/\lambda\} = 1/\alpha^{.5}$. The CV of the posterior Gamma $= 1/\alpha'^{.5} = 1/227^{.5} =$ **.066**.

Solution 5.3.5: The Prior Gamma has parameters $\alpha = 3$ and $\lambda = 4$. The Posterior Gamma has parameters $(3 + 1)$, $(4 + 2)$ $= 4,6$: $f(\theta) = (6^4/(3!))e^{-6\theta}\theta^3 = 216e^{-6\theta}\theta^3$. Thus the posterior chance that the Poisson parameter, θ, is between 1 and 2 is the integral from 1 to 2 of $f(\theta)$: $\int_1^2 216e^{-6\theta}\theta^3 d\theta$.

Solution 5.3.6: Prior Gamma has (inverse) scale parameter $\lambda = 1000$ and shape parameter $\alpha = 150$. After the first year of observations: the new (inverse) scale parameter $\lambda' = \lambda +$ number of exposures $= 1000 + 1500 = 2500$, and the new shape parameter $\alpha' = \alpha +$ number of claims $= 150 + 300 = 450$. Similarly, after the second year of observations: $\lambda'' = \lambda' + 2{,}500 = 5{,}000$ and $\alpha'' = \alpha' + 525 = 975$. The Bayesian estimate = the mean of the posterior Gamma $= \alpha''/\lambda'' = 975/5000 =$ **.195**.

Comment: One can go directly from the prior Gamma to the Gamma posterior of both years of observations, by just adding in the exposures and claims observed over the whole period of time. One would obtain the same result. Note that one could proceed through a sequence of many years of observations in exactly the same manner as shown in this question.

Solution 5.4.1: Using [5.4.1] the credibility weighted estimate is $(\alpha + N)/(\lambda + Y) = (2 + 4)/(8 + 4) = 6/12 =$ **.50**. For an alternative solution, we start with the Bühlmann Credibility parameter $K = \lambda = 8$. Then the credibility for four years is $Z = 4/(4 + 8) = 1/3$. So the credibility estimate is $(1/3)(4/4) + (1 - 1/3)(2/8) =$ **.50**.

Solution 5.4.2: Mean of the posterior distribution of θ is $(\alpha + N)/(\lambda + Y) = (250 + 89(1) + 4(2) + 1(3))/(2000 + 1000) =$ **.117**.

Solution 5.4.3: This is a Gamma-Poisson whose Prior Gamma has (inverse) scale parameter $\lambda = 1000$ and shape parameter $\alpha = 150$. Thus the Bühlmann Credibility Parameter is $K = \lambda = 1000$. One observes a total of $1500 + 2500 = 4000$ exposures. Therefore, $Z = 4000/(4000 + K) = .8$. The prior estimate is the mean of the prior Gamma, $\alpha/\lambda = .15$. The observed frequency is $825/4000 = .206$. Thus the new estimate is: $(.8)(.206) + (.2)(.15) =$ **.195**.

Comment: Same result as question 5.3.6. For the Gamma-Poisson, the Bühlmann Credibility estimate is equal to that from Bayesian Analysis.

Solution 5.4.4: EPV = Expected Value of the Poisson Means = *o*verall mean = 5.

Solution 5.4.5: The overall mean is 5. Second moment $= .1 \int_0^{10} x^2 \, dx = 100/3$. Therefore, VHM $= 33.333 - 5^2 =$ **8.333**.

Comment: For the uniform distribution on the interval (a,b), the Variance $= (b - a)^2/12$. In this case with $a = 0$ and $b = 10$, the variance is $100/12 = 8.333$.

Solution 5.4.6: $K = \text{EPV}/\text{VHM} = .6$. For one year, $Z = 1/(1 + .6) = 5/8$. New estimate is $(7)(5/8) + (5)(3/8) =$ **6.25**.

Solution 6.2.1: (1) $Z = 3/(3 + 5) = 3/8$. Proposed rate $= (3/8)(15) + (1 - 3/8)[(18)(.9)] =$ **\$15.75**. (2) $Z = 3/(3 + 10) = 3/13$. Proposed rate $= (3/13)(15) + (1 - 3/13)[(18)(.9)] =$ **\$15.92**. (3) $Z = \sqrt{3/60} = .224$. Proposed rate $= (.224)(15) + (1 - .224)[(18)(.9)] =$ **\$15.93**.

Comment: Note how the proposed rates vary much less than the credibility parameters and/or formulas.

Solution 6.4.1: (1) Reported experience = \$4 million/3,000 = \$1,333 per \$100 of payroll. Expected losses = (\$20)(3,000) = \$60,000. (Note that 3,000 represents the number of units of \$100 in payroll for Angel.) $Z = 60/(60 + 80) = .429$. So, future expected losses = (.429)(\$1,333) + (1 − .429)(\$20) = **\$583 per \$100 of payroll**.

(2) Reported *limited* experience = \$200,000/3,000 = \$66.67 per \$100 of payroll. Expected *limited* experience = (\$20)(3000)(.95) = \$57,000. $Z = 57/(57 + 50) = .533$. Reported limited experience loaded for excess losses = \$66.67/.95 = \$70.18. So, future expected losses = (.533)(\$70.18) + (1 − .533)(\$20) = **\$46.75 per \$100 of payroll**.

Comment: Using unlimited losses, the prediction is that Angel's future losses will be 29 times average. Using limited losses, the prediction is that Angel's future losses will be 2.3 time average. In an application to small insureds, one might cap even the estimate based on limited losses.

Solution 6.6.1: Husband's car: mean $= (30 + 33 + 26 + 31 + 30)/5 = 30$ and sample variance $= [(30 - 30)^2 + (33 - 30)^2 + (26 - 30)^2 + (31 - 30)^2 + (30 - 30)^2]/(5 - 1) = 6.5$. *Sue's sports car*: mean $= (30 + 28 + 31 + 27 + 24)/5 = 28$ and sample variance $= [(30 - 28)^2 + (28 - 28)^2 + (31 - 28)^2 + (27 - 28)^2 + (24 - 28)^2]/(5 - 1) = 7.5$. The mean time for both cars is $(30 + 28)/2 = 29$. EPV $= (6.5 + 7.5)/2 =$ **7.0**. VHM $= [(30 - 29)^2 + (28 - 29)^2]/$

$(2 - 1) - \text{EPV}/5 = 2 - 7/5 =$ **.60**. The Bühlmann K parameter is $\text{EPV}/\text{VHM} = 7.0/.60 = 35/3$. The credibility is $Z = 5/(5 + 35/3) =$ **.30**. Sue's estimated commute time in her sports car is: $(.30)(28) + (1 - .30)(29) =$ **28.7 minutes**.

Solution 6.6.2: **1.000**, **.889**, **.944**, **.867**, and **1.167**.

Solution 6.6.3: $(.5 - .75/1.000)^2 + (.5 - .5/.889)^2 + (.5 - .222/.944)^2 + (.5 - .40/.867)^2 + (.5 - .7/1.167)^2 =$ **.1480**.

Solution 6.6.4: A K value around 800 produces a sum of squared differences that is close to the minimum.

Comment: We want the modified loss ratios to be close to the overall average loss ratio. Based on this limited volume of data, a Bühlmann credibility parameter of about 800 seems like it would have worked well in the past. Actual tests would rely on much more data. This is insufficient data to enable one to distinguish between $K = 800$ and $K = 1{,}000$. Here is a graph that shows the sum of the squared differences between the modified loss ratios and the overall average loss ratio. Note that there is a range of K values that produce about the same minimum sum of squares.