

Full Credibility

 55M

According to classical credibility theory, a set of observations is said to have achieved the standard of **full credibility** when the sample size is sufficiently **large** to provide an accurate estimate without having to rely on data from a more general source. If this is the case, then $Z = 1$.

Full Credibility for Aggregate Losses

Let's revisit the All-Care Insurance Co. scenario. Based on internal observations, each Iowa insured has a total claim of 100 per year on average. If the sample size is large enough that the company can rely only on its internal data rather than both internal and external data, the observations will be given full credibility. Then, the estimate for the annual aggregate claims per insured would be

$$\begin{aligned}\text{Estimate} &= 1(100) + 0(200) \\ &= 100\end{aligned}$$

So what criteria will qualify the internal data as "large enough" for full credibility? In general, the sample is given full credibility when we are confident that the average (i.e. the sample mean) of the observations is close to the theoretical mean of the distribution. More specifically, for a reasonably small k , we want the sample mean to be within $100k\%$ of the theoretical mean. In mathematical notation, the following inequality needs to hold:

$$|\bar{S} - \mu_S| \leq k\mu_S$$

where \bar{S} and μ_S represent the sample mean and the true distribution mean of the annual aggregate claims per insured, respectively.

However, note that \bar{S} is a random variable. Since it varies based on the observations, it is unreasonable to want the relationship above to be true all of the time. Therefore, a threshold must be established: for a reasonably large p , the

sample mean is assigned full credibility if the inequality above is true **at least** $100p$ % of the time.

To reiterate, the goal is to determine a standard where the inequality below is satisfied:

$$\Pr\left(\left|\bar{S} - \mu_S\right| \leq k\mu_S\right) \geq p \quad (\text{S4.1.1.1})$$

Coach's Remarks

Do not confuse k and p . They are two different parameters to set the standard of full credibility.

- k : the amount of permitted fluctuation of the sample mean from the true mean; it is usually **small** (close to 0).
- p : the minimum probability for $\left|\bar{S} - \mu_S\right| \leq k\mu_S$ to hold; it is usually **large** (close to 1).

After determining a suitable distribution for \bar{S} and selecting an appropriate k and p , we are then ready to determine the minimum sample size needed for full credibility. This sample size is also referred to as the *number of exposures*. In addition, we also find it useful to restate this minimum requirement in terms of the *expected number of claims*.

Coach's Remarks

From this point on, the idea of a **minimum** requirement for full credibility will be implied; there will be no need to explicitly state "**minimum** number of exposures" or "**minimum** expected number of claims."

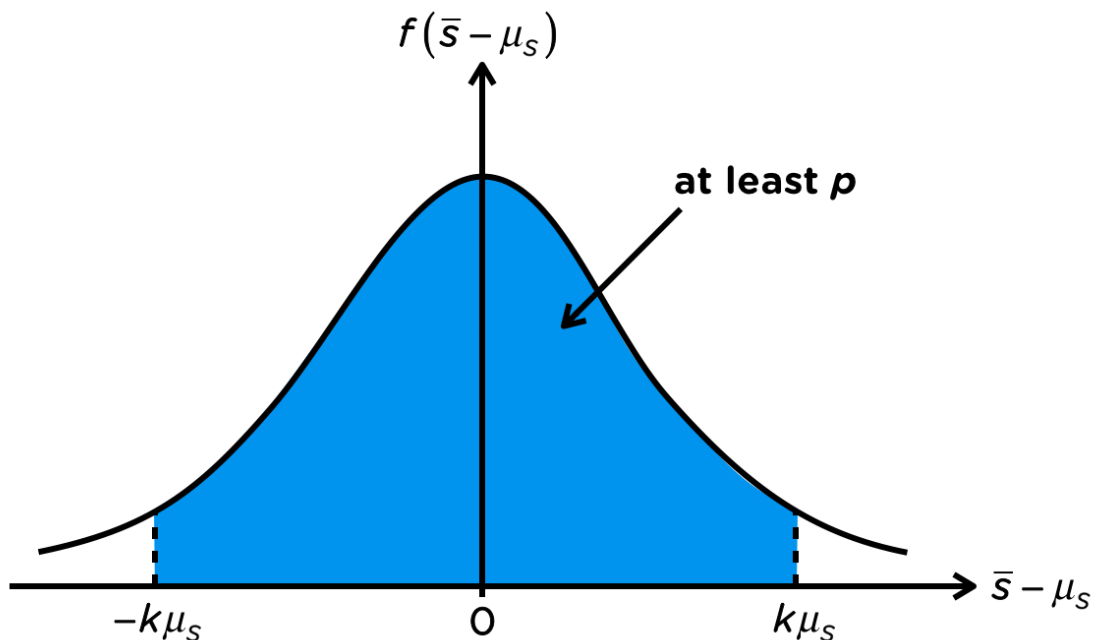
Number of Exposures

In the Credibility section, one observation/data point from the sample represents an **exposure**. Note that exposures can represent different things based on context.

For example, if All-Care wants to use one exposure to model the aggregate claims per insured per year, it will have to observe **one insured** for **one year**. If the company wants to use two exposures instead, it can either:

- observe **two insureds** for **one year**, or
- observe **one insured** for **two years**.

If the number of exposures is large, we can apply the Central Limit Theorem. Using the normal approximation, $\Pr\left(|\bar{S} - \mu_S| \leq k\mu_S\right) \geq p$ can be illustrated as:



By manipulating the inequality, the number of exposures required to assign full credibility for aggregate claims is

$$n_e = \left[\frac{z_{(1+p)/2}}{k} \right]^2 (CV_S^2) \quad (\text{S4.1.1.2})$$

where

$$\bullet \quad z_{(1+p)/2} = \Phi^{-1}\left(\frac{1+p}{2}\right)$$

- CV_S represents the coefficient of variation of the aggregate loss per exposure distribution

The proof is provided in the appendix at the end of this section.

When working with a collective risk model, (S4.1.1.2) can be expressed in terms of the claim frequency (N) and the claim severity (X):

$$\begin{aligned} n_e &= \left[\frac{z_{(1+p)/2}}{k} \right]^2 \left(\frac{\sigma_S^2}{\mu_S^2} \right) \\ &= \left[\frac{z_{(1+p)/2}}{k} \right]^2 \left(\frac{\mu_N \sigma_X^2 + \sigma_N^2 \mu_X^2}{\mu_N^2 \cdot \mu_X^2} \right) \end{aligned}$$

Let's work on the short example below using (S4.1.1.2).

You are given that the annual total claim size for each policyholder of a company follows a gamma distribution with mean 4 and variance 9.

If the full credibility standard is for the aggregate claims to be within 5% of the mean 95% of the time, what is the **number of exposures** required for full credibility?

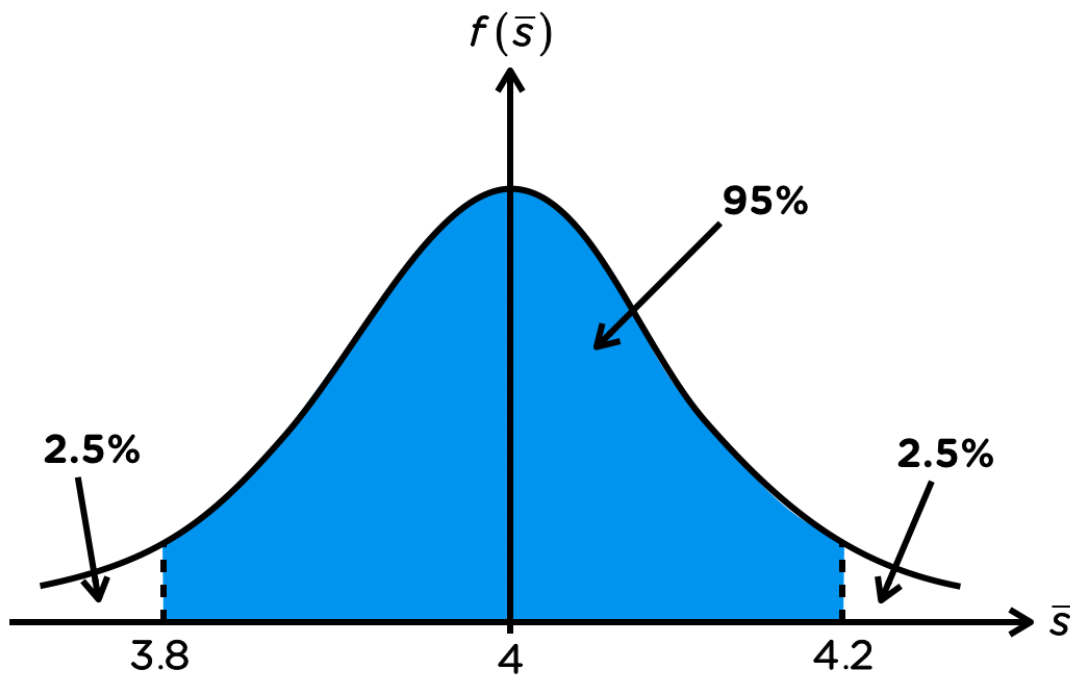
Determine each component required in the formula:

- $z_{(1+0.95)/2} = z_{0.975} = 1.95996$
- $k = 0.05$, because we want the aggregate claims to be within 5% of the mean.
- $CV_S = \frac{\sigma_S}{\mu_S} = \frac{\sqrt{9}}{4} = 0.75$

Therefore,

$$n_e = \left[\frac{1.95996}{0.05} \right]^2 \cdot 0.75^2 = 864.32$$

Graphically, if there are 864.32 exposures, then the average annual claim size will be between 3.8 and 4.2 with 95% probability.



In other words, we need at least 864.32 exposures for the sample data to be considered self-sufficient in this context. This can be achieved in many different ways; for example, we could observe 865 policyholders for 1 year or observe 173 policyholders for 5 years.

Example S4.1.1.1

As an actuary, Winson models his employer's number of claims per week using the following distribution:

Number of Claims per Week	Probability
0	0.3
1	0.25
2	0.35

Number of Claims per Week	Probability
3	0.1

It is believed that each claim has a size of 10.

Using classical credibility theory, Winson sets a standard of full credibility so that the actual aggregate claims is within 6% of the expected number 95% of the time.

Determine the number of exposures needed for full credibility.

Solution

Let N be the weekly number of claims.

$$\begin{aligned} E[N] &= 0.3(0) + 0.25(1) + 0.35(2) + 0.1(3) \\ &= 1.25 \end{aligned}$$

$$\begin{aligned} E[N^2] &= 0.3(0)^2 + 0.25(1)^2 + 0.35(2)^2 + 0.1(3)^2 \\ &= 2.55 \end{aligned}$$

$$\begin{aligned} \text{Var}[N] &= 2.55 - 1.25^2 \\ &= 0.9875 \end{aligned}$$

Next, calculate the mean and variance of the aggregate claims per week.

$$\begin{aligned} E[S] &= E[N \cdot 10] \\ &= 10(1.25) \\ &= 12.5 \end{aligned}$$

$$\begin{aligned} \text{Var}[S] &= \text{Var}[N \cdot 10] \\ &= 10^2(0.9875) \\ &= 98.75 \end{aligned}$$

Using (S4.1.1.2),

- $z_{(1+0.95)/2} = z_{0.975} = 1.95996$
- $k = 0.06$
- $CV_S = \frac{\sqrt{\text{Var}[S]}}{E[S]} = \frac{\sqrt{98.75}}{12.5} = 0.795$

Therefore,

$$n_e = \left[\frac{1.95996}{0.06} \right]^2 \cdot 0.795^2 = 674.387$$

Example S4.1.1.2

You are given:

- The annual number of claims follows a Poisson distribution with mean 5.
- The individual claim severity follows a gamma distribution with $\alpha = 3$ and $\theta = 10$.
- The claim frequency and severity are independent.
- The aggregate loss has to be within 3% of the expected loss with a probability of 90%.

Using limited fluctuation credibility theory, determine the number of exposures needed for full credibility.

Solution

Let N and X be the annual number of claims and the loss severity,

respectively.

$$N \sim \text{Poisson}(5)$$

$$X \sim \text{Gamma}(3, 10)$$

Since N and X are independent, apply the compound mean and variance formulas, i.e., (S2.5.2.1) and (S2.5.2.2), to calculate the mean and variance of the aggregate loss per year, S .

$$\begin{aligned}\mu_S &= \mu_N \cdot \mu_X \\ &= 5 \cdot (3 \cdot 10) \\ &= 150\end{aligned}$$

$$\begin{aligned}\sigma_S^2 &= \mu_N \sigma_X^2 + \sigma_N^2 \mu_X^2 \\ &= 5(3 \cdot 10^2) + 5(3 \cdot 10)^2 \\ &= 6,000\end{aligned}$$

$$CV_S = \frac{\sqrt{6,000}}{150} = 0.5164$$

Using (S4.1.1.2),

- $z_{(1+0.90)/2} = z_{0.95} = 1.64485$
- $k = 0.03$
- $CV_S = 0.5164$

$$n_e = \left[\frac{1.64485}{0.03} \right]^2 \cdot 0.5164^2 = 801.639$$

Expected Number of Claims

Besides using the number of exposures, the standard of full credibility can also be expressed in terms of the expected number of claims.

Let's rework the short example above, but this time, we add an additional piece of information and tweak the question.

You are given that the annual total claim size for each policyholder of a company follows a gamma distribution with mean 4 and variance 9.

The annual claim frequency per policyholder has a mean of 5.

If the full credibility standard is for the aggregate claims to be within 5% of the mean 95% of the time, what is the **expected number of claims** required for full credibility?

We previously calculated the number of exposures needed for full credibility to be:

$$n_e = 864.32$$

To have 864.32 exposures, we can observe the total claims of 864.32 insureds for one year. Since every insured has an annual mean of 5 claims, the expected number of claims needed for full credibility, denoted by n_c , is simply:

$$n_c = 864.32 \cdot 5 = \mathbf{4,321.62}$$

In other words, we need at least 4,321.62 (expected) claims from the sample to say that it is self-sufficient in this context.

Therefore, in general,

$$n_c = n_e \cdot \mu_N \quad (\text{S4.1.1.3})$$

where μ_N is the mean frequency per exposure.

Coach's Remarks

Do not confuse the mean frequency per exposure, μ_N , with the expected number of claims required for full credibility, n_c . The former is the frequency mean of **each** exposure, while the latter is the expected number of claims needed by the **entire** sample to meet the full credibility standard.

Instead of calculating the expected number of claims, n_c , via the number of exposures, n_e , we could combine (S4.1.1.2) and (S4.1.1.3) to calculate it directly. In general,

$$\begin{aligned} n_c &= \left[\frac{z_{(1+p)/2}}{k} \right]^2 \left(\frac{\sigma_N^2}{\mu_N} + CV_X^2 \right) \\ &= \left[\frac{z_{(1+p)/2}}{k} \right]^2 \left(\frac{\sigma_N^2}{\mu_N} + \frac{\sigma_X^2}{\mu_X^2} \right) \end{aligned} \quad (\text{S4.1.1.4})$$

where N and X are independent random variables that denote the claim frequency per exposure and the claim severity, respectively.

Full Credibility for Frequency or Severity

So far, we've discussed the full credibility standard based on aggregate claims, i.e. the standard set by:

$$\Pr\left(\left| \bar{S} - \mu_S \right| \leq k\mu_S\right) \geq p$$

However, it is possible for a full credibility standard to be based on either frequency or severity instead. Let's be clear on what this means:

- Based on frequency
 - Full credibility is met by satisfying $\Pr\left(|\bar{N} - \mu_N| \leq k\mu_N\right) \geq p$
- Based on severity
 - Full credibility is met by satisfying $\Pr\left(|\bar{X} - \mu_X| \leq k\mu_X\right) \geq p$,
but now assume that frequency is fixed, i.e. non-random

In either case, it turns out that we can calculate the standard for full credibility in terms of n_c by simply modifying (S4.1.1.4).

- If the standard is based on **frequency only**, set the severity component to 0.

$$n_c = \left[\frac{z_{(1+p)/2}}{k} \right]^2 \left(\frac{\sigma_N^2}{\mu_N} + 0 \right)$$

- If the standard is based on **severity only**, set the frequency component to 0.

$$n_c = \left[\frac{z_{(1+p)/2}}{k} \right]^2 (0 + CV_X^2)$$

Then, to express the standard for full credibility in terms of n_e instead, rearrange (S4.1.1.3).

$$n_c = n_e \cdot \mu_N \quad \Leftrightarrow \quad n_e = \frac{n_c}{\mu_N}$$

Coach's Remarks

While n_c represents the **expected** number of claims needed for full credibility,

sometimes it is okay to drop the term "expected," and call it the "number of claims needed for full credibility." This is because in practice, the observed claim frequency can be used to decide if the standard of full credibility is met.

For example, if the expected number of claims needed for full credibility is 200, then we may assign full credibility to the observations when we have at least 200 observed claims.

Thus, be aware that the word "expected" may not always be included.

Coach's Remarks

The table below shows all of the possible variations of (S4.1.1.4) for different full credibility standards and the corresponding expressions for n_e using (S4.1.1.3).

Full Credibility for ...	n_c	n_e
Aggregate Claims	$\left[\frac{z_{(1+p)/2}}{k} \right]^2 \left(\frac{\sigma_N^2}{\mu_N} + CV_X^2 \right)$	$\left[\frac{z_{(1+p)/2}}{k} \right]^2 (CV_S^2)$
Claim Frequency	$\left[\frac{z_{(1+p)/2}}{k} \right]^2 \left(\frac{\sigma_N^2}{\mu_N} \right)$	$\left[\frac{z_{(1+p)/2}}{k} \right]^2 (CV_N^2)$
Claim Severity	$\left[\frac{z_{(1+p)/2}}{k} \right]^2 (CV_X^2)$	$\left[\frac{z_{(1+p)/2}}{k} \right]^2 (CV_X^2) \frac{1}{\mu_N}$

However, keep in mind that not all formulas in this table should be memorized. They can be derived from modifying (S4.1.1.4) and applying (S4.1.1.3).

In general, consider the following strategy when approaching a full credibility question:

1. Identify the basis for full credibility: aggregate claims, frequency, or severity.
2. To find n_c , the expected number of claims needed for full credibility, use (S4.1.1.4), modifying it appropriately based on Step 1.

3. To find n_e , the number of exposures needed for full credibility, first calculate n_c , then divide it by μ_N . However, if full credibility is for aggregate claims, consider using (S4.1.1.2) instead.

Let's see this strategy in action.

Example S4.1.1.3

You are given:

- The number of claims per year follows a Poisson distribution.
- The claim amount distribution has a mean of 100 and a variance of 600.
- The number of claims and claim sizes are independent.

Determine the expected number of claims needed for full credibility if it is required that

1. the **number of claims** is within 2% of its estimate with a 95% chance.
2. the **aggregate claim** is within 8% of its estimate with a 99% chance.

Solution to (1)

Let N be the annual number of claims. Since it follows a Poisson distribution, we know that

$$\mu_N = \sigma_N^2$$

To calculate the expected number of claims needed for full credibility, apply (S4.1.1.4).

$$n_c = \left[\frac{z_{(1+p)/2}}{k} \right]^2 \left(\frac{\sigma_N^2}{\mu_N} + CV_X^2 \right)$$

Determine each component:

- $z_{(1+0.95)/2} = z_{0.975} = 1.95996$
- $k = 0.02$
- $\frac{\sigma_N^2}{\mu_N} = 1$ because $\mu_N = \sigma_N^2$.
- $CV_X^2 = 0$ because the full credibility is calculated for the **claim frequency only**.

Therefore,

$$n_c = \left[\frac{1.95996}{0.02} \right]^2 (1 + 0) = \mathbf{9,604}$$

Solution to (2)

Let N be the annual number of claims and X be the claim amount.

Again, to calculate the expected number of claims needed for full credibility, apply (S4.1.1.4) and determine each component:

- $z_{(1+0.99)/2} = z_{0.995} = 2.57583$
- $k = 0.08$
- $\frac{\sigma_N^2}{\mu_N} = 1$ because $\mu_N = \sigma_N^2$.
- $CV_X^2 = \frac{\sigma_X^2}{\mu_X^2} = \frac{600}{100^2} = 0.06$

Therefore,

$$n_c = \left[\frac{2.57583}{0.08} \right]^2 (1 + 0.06) = \mathbf{1,099}$$

Example S4.1.1.4

You are given:

- The annual number of claims has the following distribution:

Number of Claims	Probability
0	0.3
1	0.4
2	0.1
3	0.2

- The claim amount follows a distribution with the following density function:

$$f(x) = \frac{1}{x\sqrt{4\pi}} \exp\left(-\frac{[\ln x - 5]^2}{4}\right), \quad x > 0$$

- The number of claims and claim sizes are independent.

The credibility standard is set so that the aggregate claim is within 5% of the mean 90% of the time.

Calculate

- the **expected number of claims** needed to achieve the full credibility standard.
- the **number of exposures** needed to achieve the full credibility standard.

Solution to (1)

Let N and X be the annual number of claims and the loss severity, respectively.

To calculate the expected number of claims needed for full credibility, apply (S4.1.1.4).

$$n_c = \left[\frac{z_{(1+p)/2}}{k} \right]^2 \left(\frac{\sigma_N^2}{\mu_N} + CV_X^2 \right)$$

Determine each component:

- $z_{(1+0.90)/2} = z_{0.95} = 1.64485$
- $k = 0.05$
- For the frequency distribution,

$$\begin{aligned} E[N] &= 0.3(0) + 0.4(1) + 0.1(2) + 0.2(3) \\ &= 1.2 \end{aligned}$$

$$\begin{aligned} E[N^2] &= 0.3(0^2) + 0.4(1^2) + 0.1(2^2) + 0.2(3^2) \\ &= 2.6 \end{aligned}$$

$$\text{Var}[N] = 2.6 - 1.2^2 = 1.16$$

- For the severity distribution, we can recognize from the density function that it is lognormal with $\mu = 5$ and $\sigma^2 = 2$.

$$X \sim \text{Lognormal}(5, 2)$$

Using the exam table,

$$E[X] = e^{5+0.5(2)} = 403.43$$

$$E[X^2] = e^{2(5)+2(2)} = 1,202,604.28$$

$$\begin{aligned} \text{Var}[X] &= 1,202,604.28 - 403.43^2 \\ &= 1,039,849.49 \end{aligned}$$

Therefore,

$$\begin{aligned} n_c &= \left[\frac{1.64485}{0.05} \right]^2 \left(\frac{1.16}{1.2} + \frac{1,039,849.49}{403.43^2} \right) \\ &= \mathbf{7,960.46} \end{aligned}$$

Solution to (2)

To calculate the number of exposures needed for full credibility, apply (S4.1.1.3).

$$n_c = n_e \cdot \mu_N$$

From Solution to (1), we have $n_c = 7,961.91$ and $\mu_N = 1.2$.

Therefore,

$$7,960.46 = n_e \cdot 1.2$$

$$n_e = \mathbf{6,633.71}$$

Alternative Solution to (2)

Alternatively, (S4.1.1.2) can be applied to answer this question. Since N and X are independent, we can use the compound mean and variance formulas to calculate CV_S , where S is the aggregate claims per year.

$$\begin{aligned} E[S] &= E[N]E[X] \\ &= 1.2 \cdot 403.43 \\ &= 484.11 \end{aligned}$$

$$\begin{aligned} \text{Var}[S] &= E[N]\text{Var}[X] + \text{Var}[N]E[X]^2 \\ &= 1.2(1,039,849.49) + 1.16(403.43)^2 \\ &= 1,436,614.95 \end{aligned}$$

$$CV_S = \frac{\sqrt{1,436,614.95}}{484.11} = 2.476$$

$$\begin{aligned} n_e &= \left[\frac{1.64485}{0.05} \right]^2 \cdot 2.476^2 \\ &= \mathbf{6,633.71} \end{aligned}$$

Example S4.1.1.5

For an insurance portfolio, you are given:

- The annual claim counts follow a negative binomial distribution with parameters $r = t$ and $\beta = 5$.
- The claim sizes follow a gamma distribution with $\alpha = 4$ and an unknown constant θ .

The annual claim counts and claim sizes are independent.

Using classical credibility, 1,691 expected claims are needed so that the aggregate claim is within 100 k % of the expected value with 90% probability.

Determine k .

Solution

Let N and X denote the annual claim count and claim size, respectively.

From the question,

- $N \sim \text{Negative Binomial}(t, 5)$
- $X \sim \text{Gamma}(4, \theta)$

To calculate the expected number of claims needed for full credibility, apply (S4.1.1.4).

$$n_c = \left[\frac{z_{(1+p)/2}}{k} \right]^2 \left(\frac{\sigma_N^2}{\mu_N} + CV_X^2 \right)$$

Determine each component:

- $n_c = 1,691$
- $z_{(1+0.90)/2} = z_{0.95} = 1.64485$
- For the frequency distribution,

$$\mu_N = 5t$$

$$\sigma_N^2 = 5t(1 + 5) = 30t$$

- For the severity distribution,

$$\mu_X = 4\theta$$

$$\sigma_X^2 = 4\theta^2$$

Therefore,

$$\begin{aligned} 1,691 &= \left[\frac{1.64485}{k} \right]^2 \left(\frac{30t}{5t} + \frac{4\theta^2}{[4\theta]^2} \right) \\ &= \left[\frac{1.64485}{k} \right]^2 \left(6 + \frac{1}{4} \right) \end{aligned}$$

$$\left(\frac{1.64485}{k} \right)^2 = 270.56$$

$$k = 0.1$$

Example S4.1.1.6

You are given the following:

- Annual number of claims for each insured follows a binomial distribution with parameters m and $q = 0.2$.
- Claim severity is independent of the annual claim frequency per insured, and has a coefficient of variation λ .
- Using classical credibility theory, 3,000 claims are needed to achieve the standard of full credibility.

Determine the value of λ such that

1. the full credibility standard is set so that the claim severity is within 5% of its mean 95% of the time.
2. the full credibility standard is set so that the pure premium is within 5% of its mean 95% of the time.

Solution to (1)

To calculate the expected number of claims needed for full credibility, apply (S4.1.1.4).

$$n_c = \left[\frac{z_{(1+p)/2}}{k} \right]^2 \left(\frac{\sigma_N^2}{\mu_N} + CV_X^2 \right)$$

Determine each component:

- $n_c = 3,000$
- $z_{(1+0.95)/2} = z_{0.975} = 1.95996$
- $k = 0.05$
- $\frac{\sigma_N^2}{\mu_N} = 0$ because the full credibility is calculated for the **claim severity only**.
- $CV_X^2 = \frac{\sigma_X^2}{\mu_X^2} = \lambda^2$

Therefore,

$$\begin{aligned} 3,000 &= \left[\frac{1.95996}{0.05} \right]^2 (0 + \lambda^2) \\ &= 1,536.58 \cdot \lambda^2 \end{aligned}$$

$$\begin{aligned} \lambda &= \sqrt{\frac{3,000}{1,536.58}} \\ &= \mathbf{1.397} \end{aligned}$$

Solution to (2)

The term "**pure premium**" is just another way to represent the aggregate claims per exposure. That means we can apply the formulas that pertain to full credibility for aggregate claims.

Let N be the annual number of claims per insured.

$$N \sim \text{Binomial}(m, 0.2)$$

To calculate the expected number of claims needed for full credibility, apply (S4.1.1.4).

$$n_c = \left[\frac{z_{(1+p)/2}}{k} \right]^2 \left(\frac{\sigma_N^2}{\mu_N} + CV_X^2 \right)$$

Determine each component:

- $n_c = 3,000$
- $z_{(1+0.95)/2} = z_{0.975} = 1.95996$
- $k = 0.05$
- $\frac{\sigma_N^2}{\mu_N} = \frac{m(0.2)(0.8)}{m(0.2)} = 0.8$
- $CV_X^2 = \frac{\sigma_X^2}{\mu_X^2} = \lambda^2$

Therefore,

$$\begin{aligned} 3,000 &= \left[\frac{1.95996}{0.05} \right]^2 (0.8 + \lambda^2) \\ &= 1,536.64(0.8 + \lambda^2) \end{aligned}$$

$$0.8 + \lambda^2 = 1.952$$

$$\begin{aligned}\lambda &= \sqrt{1.952 - 0.8} \\ &= \mathbf{1.073}\end{aligned}$$

Coach's Remarks

Most classical credibility questions have very similar but unintuitive wordings. Therefore, it is important to know how to extract information from these questions.

Consider the following quotes from a few SOA sample problems.

- **FAM-S Sample Q1:** "The observed pure premium should be within 2% of the expected pure premium 90% of the time."