Canadian Institute of Actuaries / Institut canadien des actuaires

SOCIETY OF ACTUARIES®

The Application of Credibility Theory in the Canadian Life Insurance Industry

November 2019

The Application of Credibility Theory in the Canadian Life Insurance Industry

AUTHORS

Leslie M. Jones, ASA, MAAA
Consulting Actuary
Risk & Regulatory Consulting

Patricia Matson, FSA, MAAA
Partner
Risk & Regulatory Consulting

SPONSORS

Canadian Institute of Actuaries (CIA)
Society of Actuaries (SOA)

Contents

The Application of Credibility Theory in the Canadian Life Insurance Industry

Executive Summary

The Canadian Institute of Actuaries (CIA) and the Society of Actuaries (SOA) engaged Risk & Regulatory Consulting, LLC to perform a study on the application of credibility theory in the Canadian life insurance industry (the Study). The Study was composed of two components: a survey component and an analysis component. Canadian life and annuity companies[1] were surveyed in order to gain an understanding of the approaches used to assess data credibility. An analysis and comparison of the credibility methods being utilized was performed for both mortality and lapse using representative sample data sets (data constructed by the researchers showing representative industry mortality and lapse experience). A summary of the key findings follows:

Key Survey Findings

1. Most of the companies surveyed stated they use the Limited Fluctuation Credibility Theory (LFCT) method to determine credibility for mortality. Some of the responding companies indicated they use LFCT for lapse, with the remainder not using any form of formal credibility method for lapse. Key drivers behind the choice of LFCT were reported to be the availability of CIA guidelines and the simplicity of the method.
2. Based on the survey responses, no other formal credibility methods are currently being used. Companies that did not use a formal credibility method generally reported using 100% industry data or 100% company data with ad-hoc adjustments.
3. Most of the companies that reported using LFCT indicated they use the “by number” approach (i.e., the Simple Poisson Model, also referred to by the researchers as “by policy” or “by count”) to calculate credibility factors. Several companies reported using 3,007 deaths as the criterion for full credibility.[2]
4. A limited number of companies reported using the “by amount” approach (i.e., the Compound Poisson Model) to reflect the financial impact of the assumptions and to capture variability of claim size to better reflect the company’s exposure.
5. Other industry methods, such as the Bayesian Credibility and the Greatest Accuracy Credibility Theory (GACT, also known as *Bühlmann Credibility* and *Linear Bayesian Credibility*), have been considered by some companies but ultimately rejected because of their complexity.

Key Analysis Findings[3]

1. Expected results based on theoretical aspects of LFCT and the nature of the sample data sets:
    a. As expected, the number of occurrences (deaths or lapses) needed for full credibility “by count” do not vary by assumption (mortality or lapse), product type or risk classification. This is due to the nature of the Simple Poisson Model which is used to develop the results for the LFCT “by count.” The only parameters which impact the number of occurrences needed for full credibility for the Simple Poisson Model are the confidence level and margin of error selected. Consequently, determining the number of occurrences needed for full credibility is directly related to the level of

[1] The reference to “Canadian life and annuity companies” is a reference to the type of company included in the Study. However, the focus of the Study was individual life insurance.
[2] The consensus recommendation of the CIA Educational Note (see below) for the minimum number of deaths needed for 100% (i.e., “full”) credibility is 3,007. CIA Educational Note, page 16, item 3.
[3] Early in the project, the researchers learned that there is really only one credibility method typically used in the Canadian industry, namely the LFCT. Therefore, the research and analysis is focused on that method.

precision sought in establishing assumptions. This result is discussed further in Appendix C and in the CIA Educational Note.[4]

   b. The number of occurrences (deaths or lapses) needed for full credibility "by amount" are materially higher than the number needed for full credibility "by count" for both mortality and lapse for all of the product types and risk classifications analyzed with the sample data sets. This result is not unexpected given that the underlying exposure data for the sample data sets is the CIA mortality data, and as such there is significant dispersion in the exposure in the blocks under consideration.[5]

2. The number of occurrences (deaths or lapses) needed for full credibility "by count" and "by amount" both vary depending on the confidence level and the margin of error selected. Decreasing the margin of error and increasing the confidence level both result in an increase in the number of occurrences needed for full credibility. However, decreasing the margin of error results in a much larger increase in the number of occurrences needed for full credibility than increasing the confidence level both "by count" and "by amount" for all product types and risk classifications for the sample data sets.
3. The number of occurrences (deaths or lapses) needed for full credibility "by amount" varies by product type and risk classification for the sample data. These results are driven in part by the relative variation in the exposure of these data sets. The number of occurrences needed for full credibility "by amount" varies depending on the characteristics of the underlying portfolio and the assumptions used.
4. The number of occurrences (deaths or lapses) needed for full credibility 'by amount" may depend on factors other than the dispersion in the underlying exposure data. In our analysis, the number of lapses needed for full credibility "by amount" is higher than the number of deaths needed for full credibility "by amount" for one of the products analyzed in the sample data sets. The result is the opposite for the other product analyzed. These results are driven at least in part by the difference in the mortality rates and the lapse rates for the sample data sets. The data sets were constructed so that the underlying exposures would be the same for comparison purposes.
5. The researchers note that aggregating exposures by duration or by attained age may significantly reduce the indicated number of occurrences (lapses or deaths) needed for full credibility "by amount." This simplified approach masks some of the variability in the underlying data, because it assumes that the exposure and expected decrement are the same for all lives at a particular attained age or duration.

These results are based on the sample data sets constructed by the researchers. Actual results for a company will depend on the nature of the company's block of business, including the amount of variation in net amount at risk for the company's block.

[4] CIA Educational Note, page 16, items 2 and 3.
[5] CIA Educational Note, page 16, item 2.

Section 1: Introduction

1.1 An Introduction to the Study

The Canadian Institute of Actuaries (CIA) and the Society of Actuaries (SOA) engaged Risk & Regulatory Consulting, LLC ("the researchers") to perform a study on the application of credibility theory in the Canadian life insurance industry and summarize results in a paper to be published by the CIA and SOA. The CIA and SOA established a project oversight group (POG) to work closely with the researchers to carry out the main objectives of the project. The Study consisted of two parts:

1. The first part involved a survey of Canadian life and annuity companies on credibility theory practices. This survey explored credibility methods being used, the associated processes employed and data used.
2. The second part involved comparing and contrasting the methods that are being utilized using sample data sets for both mortality and lapse.

The purpose of the Study is to create reference material to assist practicing life actuaries and other practitioners in understanding and applying credibility theory. Early in the project, the researchers learned that there is really only one credibility method typically used in the Canadian industry, namely Limited Fluctuation Credibility Theory (LFCT). Therefore, the research and analysis is focused on that method.

Based on the results of the Study analysis, we have summarized the approach, information gathered and conclusions and shared the Study results with the POG for additional input and feedback. The summary includes information that is responsive to each of the objectives outlined above.

This report presents the methodological approach applied in the Study, and the primary results of the Study. Survey results are summarized from the research conducted, and do not represent the views or opinions of the researchers or the POG.

The Study did not consider the impact that the implementation of IFRS 17 may have on the general approach for setting assumptions, including any related credibility theory practices.[6]

The researchers and the POG note the following as potential topics for further exploration:

1. The impact that the implementation of IFRS 17 may have on the general approach for setting assumptions, including any related credibility theory practices.
2. Further exploration of the appropriate application of the LFCT in setting assumptions, including additional study of the LFCT "by amount" and application of the LFCT method to the lapse assumption.
3. Alternative methods to use for determining the credibility of small numbers of observations. For example, the situation where there is a significant amount of exposure but limited occurrences of the event of interest.
4. Approaches to making other more complex credibility methods, such as the Greatest Accuracy Credibility Theory (GACT), more accessible to companies. This may involve a study developed from actual company data that is collected for recurring CIA mortality and lapse studies, with any needed additional fields added to the data collection to enable the analysis of alternative credibility methods (e.g., GACT).
5. Further exploration of alternative credibility methods (i.e., alternative to LFCT and GACT) that may be appropriate for risks born by the life insurance industry.

[6] Readers are encouraged to communicate issues or needs where credibility may be a concern in the context of IFRS 17 to the CIA IFRS 17 Steering Committee. Les Rehbeli, FCIA, FSA, MAAA, was the point of contact for life and health for that group during the writing of this paper.

1.2 An Introduction to Credibility Theory

The application of credibility theory is often required to evaluate the appropriateness of assumptions such as mortality and lapse levels for a company's block of business. To successfully apply the theory, an actuary needs a good understanding of available credibility methods and their uses and limitations.

A company's own experience for a particular block of data is usually the most relevant source of data.[7] Thus, in an ideal setting a company would be able to rely entirely on its own experience studies to establish assumptions. However, in many instances company experience may not be available or sufficient to adequately establish assumptions. In these cases, the company may need to rely on external sources of data or judgment to establish assumptions. Credibility theory may be used to help a company assess whether or not its data is "fully credible" or "100% credible," in which case companies may develop assumptions or create tables based on their own data. If the data is not fully credible, then credibility theory methods may be used to combine the company experience with appropriate base experience (e.g., an industry table or a prescribed valuation table) to develop a more accurate estimate. It is important to note that if appropriate base experience is not available or credible, it may be necessary to rely on other sources of information or actuarial judgment rather than applying credibility theory to partially credible data. Also, there are a variety of factors to consider in assembling or adjusting the data for the company's experience and the selection of the base experience that are beyond the scope of this analysis. However, the reader is referred to the CIA Educational Note and the other documents referenced in Appendix B for more information on these important considerations.

Once available, company experience data (which may not be fully credible) and appropriate base experience (which is assumed to be fully credible) have been suitably prepared and segmented, they may be blended using credibility weightings.[8]

There are two main approaches to determining the credibility weightings: GACT and LFCT. Both of these approaches are discussed in more detail in Section 4.2 and Appendix C of this report. As described above, the focus of our analysis is on LFCT.

Section 2: Survey Development

2.1 General Background

The researchers developed survey questions designed to gain an understanding of credibility theory practices of Canadian life and annuity companies. The survey questions were designed to address various topics, including:

- How are life and annuity companies applying credibility theory for mortality and lapse?
- Are life and annuity companies applying credibility theory to determine other assumptions?
- What are the internal and external guidelines for when a life and annuity company applies credibility theory?
- What are the methods and data being used and how do they vary by use? For example, are the methods and data the same for both mortality and lapse?
- What software is being used?
- How does the application of credibility theory vary by product type?

The survey questions were discussed with the POG and modified based on the input received. A complete list of survey questions is included in Appendix A.

---

[7] CIA Educational Note, Section 210. The Educational Note limits this statement to mortality experience, but the researchers believe that this statement is more broadly applicable to other company experience.
[8] CIA Educational Note, Section 500.

2.2 Selection of Survey Participants and Response Rate

The researchers worked with the POG to develop a list of Canadian life and annuity companies and to identify appropriate contacts within each company. Fifteen companies were identified through this process. These companies represented more than 95% of 2017 total market premiums in Canada. Surveys were sent to the identified contacts. The researchers received responses from 11 companies, representing approximately 70% of 2017 total market premium in Canada. Eight companies completed the survey, two provided responses to the survey in a phone interview in lieu of completing the survey and one provided a very brief e-mail response. Survey responses were collected between April and June of 2018.

The survey participants are listed in Section 5. There was no data specifically collected for our analysis. Instead, we used data from separate CIA studies previously published (see Section 4.3). Consequently, participants in these reference CIA studies have indirectly contributed to our analysis.

Section 3: Survey Results[9]

3.1 Application of Credibility Theory for Mortality and Lapse Assumption Setting

The survey responses summarized in this section did not distinguish between mortality and lapse assumption setting unless specifically noted otherwise. In addition, because some companies used more than one method, the total number of methods described below is greater than 11.

Nine out of 11 companies surveyed use the LFCT method to determine credibility for mortality. Four of these companies noted that they use the "Normalized Method" to calculate credibility factors by sub-category, which involves use of actual to expected ratios of specific sub-categories of data, rather than just total company ratios. These methods are outlined in the CIA Educational Note (Document 202037) "Expected Mortality: Fully Underwritten Canadian Individual Life Insurance Policies" published in 2002 ("CIA Guidance" or "CIA Educational Note").

Four companies reported using 3,007 deaths as the criterion for full credibility, and one company reported using 6,014 deaths. The company using 6,014 deaths explained that using 3,007 deaths assumes homogeneous policies and may understate true full credibility. The company stated that they use the CIA guidance note approach using 6,014 deaths as full credibility in order to recognize variability in coverage amounts and age distribution. Other companies echoed the concern that the 3,007 number would not be adequate when comparing non-homogenous lives. One company reported using 3,007 events for full credibility if there is a recognized external industry study. If there is not a recognized external industry study, the company varies the margin of error used in determining the criterion for full credibility based on the homogeneity of the data.

One company reported that it does not apply a credibility method because of concerns that the standards for full credibility are not high enough, and specifically that the LFCT method assigns too much weight to company data for small numbers of deaths. This company stated that it uses the LFCT method as a reference only and that it applies "100% industry for mortality and 100% company for lapse."

Although the company which reported using 6,014 deaths indicated that it uses the CIA Guidance, the researchers did not count this company as one of the nine companies using LFCT, since the company did not elaborate on the method used to derive the 6,014 deaths.

Lapse Methods

Many companies report they do not apply credibility theory for lapse assumptions, but they do validate their company experience against industry studies. Regarding lapse, the methods applied varied among the following approaches:

- LFCT.

[9] Section 3 presents a summary of the survey results. The researchers have endeavored to incorporate verbatim responses to the extent practical. However, in some instances the responses have been modified for clarity. In other instances, the researchers have noted their interpretation of the response or provided additional information about the response.

- LFCT, except when lapse rates are very high (no additional information regarding the method used when the lapse rates are very high was supplied).
- “By number” cell-based criteria (no additional information regarding the cell-based criteria was supplied; however, the researchers assume that the cells include policies or contracts that are considered to be alike with regard to various characteristics (e.g., duration, policy type, demographic characteristics, risk classification)).
- Internal lapse studies in conjunction with in-depth analysis of industry data. Typically, the company has found the internal results to be consistent with the industry data and is therefore confident in using its own results.
- Using the company’s own experience plus industry results when available, but no formal credibility approach.
- 100% credibility assigned to the company’s internal studies, with some adjustments.

Drivers of Choosing LFCT Method

The key drivers behind the choice of the LFCT method were reported to be:

- Availability of CIA guidelines.
- The simplicity of the method.
- Number of events (the researchers interpreted this response to mean that the number of events needed for full credibility was reasonable given other responses provided by the company).
- The method allows for the calculation of credibility factors by sub-category.
- The perception that this method is the preferred approach in the Canadian industry.
- For lapse, one company found this method to be the most representative of the company’s experience and view on tail risks. The company added that the method allows it to recognize more rapidly its recent experience as experience emerges (the researchers interpret this to mean that the company believes that the method assigns an appropriate level of credibility to emerging company experience).
- One company has historically used this method and does not have enough internal resources to research and test new methods, especially given that experience has been stable using this method.

Drawbacks and Issues of LFCT

The main drawback noted for the LFCT method was that it does not have a strong theoretical base. Other industry methods, such as the Bayesian Credibility and the GACT, have been considered by some companies but ultimately rejected because of their complexity. The GACT method was also reported to be avoided because it requires data of several companies and does not work well if the random variable has a heavy tail.

Some of the issues encountered, according to survey respondents, when applying the LFCT method are described herein:

- For some sub-categories, industry data is far from fully credible, requiring more weight to be applied to the company’s internal data.
- In other cases, sub-populations are low in credibility, requiring normalization using a bigger population, like industry experience. Alternatively, partial credibility can be applied by blending the current or external assumptions that correspond to the company’s business mix with actual experience.
- Past experience, even if it is credible, may not be representative of the future due to changes in business mix, product design, underwriting or the economic environment. The solution in this case is to adjust historical experience to better reflect more recent conditions.
- Too much weight on industry mortality experience (due to low credibility of the company’s internal data) could lead to more variability as experience emerges.
- For lapse, there may be limited comparability to industry experience for certain products and therefore company experience cannot be blended with industry experience.
- Applying credibility factors can lead to variability in experience which must be anticipated for budgeting and pricing purposes (implicit margins for adverse deviation).

- When applying no credibility factors and industry data is available, the company stated that it must demonstrate that its view is at least as conservative as industry studies.
- The LFCT method requires industry experience, which is not always available.
- Users must be careful not to overestimate accuracy of a specific number, especially when based on count. Some respondents expressed concern that the full credibility number (3,007) might need to be updated, and does not consider gender, smoker status, etc.
- The CIA Educational Note proposes 3,007 lives for full credibility, but volatility seems to indicate a much higher number would be required. The use of 3,007 assumes that these lives are homogeneous, independent and statistically ideal, which is not seen in practice. For a medium-sized company with products similar to the industry, using industry data is a better proxy. There is concern that the LFCT method gives too much credibility to smaller amounts of data.
- It is difficult to apply to new products where the company may not have a great deal of experience, even though that experience may be more representative than industry data.

Survey respondents also noted that the LFCT method requires judgment in using credibility to set assumptions in cases of limited experience/low incidence (limited number of occurrences). In situations of very low incidence rates, one company indicated that the standard LFCT method assigns very low credibility to company experience. This company noted that if the size of the block of business is very small, and it is unlikely that future studies will increase credibility, it generally looks at external references (industry studies, reinsurance studies, reinsurance quotations, similar products within the company, etc.) in order to validate the appropriateness of current assumptions. If there is insufficient experience for a new product, it may be preferable to wait until more experience emerges before updating assumptions. However, if early indications are that emerging experience deviates from the expectation and the assumption is material, judgment will be applied. In these cases, the company indicated that it may choose to use LFCT to validate the reasonableness of the current assumptions by reviewing confidence intervals rather than applying it to set assumptions. One respondent indicated that if the actual number of lapses deviates by more than one standard deviation from the expectation, this could indicate that the assumption is inappropriate. If it deviates by more than two standard deviations, the company believes this represents a strong indication that the assumption is incorrect. In such situations, the company stated that some judgment is required to determine the appropriate assumption.

Other solutions noted by survey respondents in situations of limited credibility include:

- Increasing the length of the experience study.
- Comparing the results with a prior (or similar) experience study, which may allow one to gauge reasonability of results in the current study.
- Reviewing actual-to-expected (A/E) results by calendar year within the experience study period. If A/E results are relatively stable across all calendar years, experience is likely more credible than LFCT may indicate.

Balancing Volume and Relevance of Data

When selecting the credibility method, it is necessary for companies to balance the need for sufficient data (which may require longer time periods) and recent data (shorter time periods). Various approaches to doing so noted by the respondents are described herein:

- Combining like products to create sufficient data. When this is not appropriate (as is generally true for mortality for newer products), industry data is used.
- One company indicated that assumptions based on 5-6 year studies are generally sufficient to get full credibility for mortality, and believes that industry best practice uses five years of data. Another company stated that it considers longer time periods to improve credibility when no industry data is available and could apply some adjustment to the resulting assumption to better recognize recent trends in experience. One company stated that a longer period would provide full credibility for lapses but would put less weight on current trends. Another company described the opposite approach, stating that 10-12 year studies are more useful for mortality and situations with limited experience and that more recent experience may be

more appropriate in situations where there is an emerging trend (more common for lapse). One company stated that it uses five years, each weighted equally. Another company stated that it used five years, but weighted the years using the "sum of the digits" method (the researchers interpret the "sum of the digits" method to be one that assigns more weight to more recent years).

- Application of mortality improvement adjustment in the final results.
- Starting an experience study by selecting data over a longer period in order to understand trends and relationships, then defining the appropriate study reference period to support the recommended assumptions.
- One company stated that it has not reflected varying credibility for different data periods.
- One company stated that it does not apply credibility or partial credibility to newer products.
- One company stated that it is less concerned about credibility when the blocks are smaller, because it is necessary to weigh the costs and benefits when considering materiality and data availability.

Most companies surveyed report that their credibility method does not vary by use (pricing, reserving or overall company risk assessment). However, one company only uses the credibility method for setting pricing assumptions. Another company uses the credibility method primarily for reserving. One company stated that it uses the credibility method for dividend scale management and pricing.

3.2 Application of Credibility Theory to Other Assumptions

Four out of 10 companies reported using the LFCT method for other assumptions such as morbidity, policyholder behavior, claim lags and long-term disability death and termination rates. One of these companies noted the use of an external guideline specific to the assumption and to the product. This company stated that it makes adjustments to the items included in the guidance to reflect the nature of the business. Another company reported that it does not use industry studies to apply credibility to the assumption. Instead, it uses its reinsurer's view on the assumption as a base assumption, which it cross-validates to its own internal results.

3.3 Internal and External Guidelines

All companies report using the CIA Educational Note as their guideline for applying credibility theory. Other guidelines referenced include the SOA Credibility Theory Practices Report,[10] the American Academy of Actuaries (AAA) Credibility Practice Note[11] and the AAA Group Long-Term Disability Valuation Standard – Section J.[12]

3.4 Sources of Data

Most companies report using a combination of internal experience and industry data when applying credibility theory. One company reported that all data comes from its administration systems, with variables/fields included based on the assumption that is being studied. One company noted that in addition to using internal experience and industry studies for mortality assumption setting, it uses its reinsurers' assumptions (e.g., mortality, morbidity) for certain products. One company noted that it mostly uses internal data for lapse but refers to external reports and may use the published rates from those reports (for example, if they do not have credibility for long terms, especially with newer products).

One company reported that it applies the LFCT to the A/E ratio to find the weighted average of industry and its own A/E, and A/E could be aggregated differently for mortality and lapse. For its participating mortality study, LFCT is used to blend the sub-category A/E ratios with the experience from the overall participating block.

One company reported segregating by gender and smoker status as well as duration, noting that it only uses data for legacy products where a lot of data is available. One company reported segregating by underwriting cohort. When

---

[10] See Reference 3 of Appendix B.
[11] See Reference 4 of Appendix B.
[12] See Reference 9 of Appendix B.

CIA industry data is not available, it blends company experience with the company's long-term assumption. Another company reported segregating by duration, age and product.

3.5 Software

All companies use Microsoft Excel (or similar spreadsheet software) combined with internal software, SAS or Microsoft Access to analyze the credibility of data. Three out of 10 companies also use Access to calculate expected claims and actual claims for each policy and to implement internal mortality studies. However, many companies appear to avoid this program because of data limitations. One company reported using Excel to do experience studies for its smaller lines of business, but using SQL for other life products because SQL can handle large volumes of data. Excel is preferred by one company because it is easy to use and to customize for the purpose of analysis. One company is planning to write its code in R and move to Tableau.

One company reported that it also uses an internal tool that scrubs data and calculates the exposure and the A/E results. One company reported using internal software to do all data analysis and then using Excel to apply credibility theory. Internal software allows customization to better fit the needs of the company and more control over the results. However, this approach is costlier due to the necessity of maintaining internal IT expertise and development costs.

3.6 Application of Credibility Theory by Product Type

Regarding application of credibility by product type, one company reported it only applies credibility theory on certain products. Eight out of nine other companies reported using the same credibility approach for all products, but one company noted that different product types can result in sensitivity in different areas of the assumption, and LFCT can be applied to different sub-groupings (for example, the assumption by age). One company reported credibility theory varies by product and by underwriting type because the application depends on the nature of the business and availability of relevant industry data for blending purposes.

3.7 Adjustments for New Products

Approaches for adjusting credibility methods for new products or changes in underwriting criteria varied among survey participants:

- If company data is not considered to reflect products that are sufficiently similar to the new product, use industry data.
- Use existing aggregate results, modifying as needed using any possible source.
- Seek guidance from external reinsurers and internal underwriters.
- Consider changes in product design, distribution channel, application method and underwriting criteria when setting the assumptions.

3.8 Supplementation of Experience Data with Industry Data

Companies reported various data sources and approaches to supplementing experience data with industry data:

- Use of the annual CIA Canadian Standard Ordinary Life Experience results for their mortality assumptions.
- Construction of mortality table based on CIA mortality data and the company's own data.
- Use of the experience of industry or population data in cases of limited experience and at younger and older ages.
- Even when data is fully credible, reviews of industry data for reasonableness, particularly where there is less data available.

3.9 Basis Risk Adjustment

Approaches for adjusting for basis risk (i.e., the risk that differences in populations may result in an inappropriate assumption) varied among survey participants:

- Recognize partial credibility and blend with the industry. At younger and older ages, grade to industry.

- Analyze industry data to see how they compare to the products offered and adjust as appropriate.
- For certain products, industry data is not expected to have a material impact.
- Assumption is split by distribution channel, reflecting differences in the population.
- No formal approach for adjusting for basis risk. For life insurance mortality study, there is no adjustment, as company results are blended with industry underwritten results.
- Adjustments to CIA mortality table based on the company's own data.

3.10 Weighting of Company Experience

The survey respondents noted the following approaches for weighting company experience data that is not fully credible:

- For lapse, interpolate between fully credible cells.
- Partial credibility factor is calculated as the square root of the company's claim divided by industry claim in each sub-category, thereby assigning more weight to internal experience when industry data is not fully credible (the researchers interpret "claim" in this response to be number of claims).
- Apply the Normalized Method outlined in the CIA Educational Note.
- Apply the credibility factor in computing the credibility weighting, taking into consideration the number of observed events and the criterion for full credibility. The company provided an example which indicates that the credibility weighting is based on the methods for LFCT "by count" described in the CIA Educational Note (see the Standard Normal Table – Range and Probability Parameters table on page 33).
- For valuation, a weighted average of company A/E and industry A/E is used, where the weight on company experience is the square root of (n/3,007), n being the number of claims for the company.
- For participating mortality pricing, the overall block experience is used to blend with the A/E for each sub-category if credibility is less than 100%. No industry experience is used.
- Weight using the Compound Poisson method (LFCT "by amount").

Some companies indicated that their mortality experience is fully credible. Thus, they do not weight company experience with industry experience (i.e., they assign 100% credibility to company experience).

3.11 Evaluation of the Credibility Method

The survey responses included the following approaches used to evaluate whether the results of a credibility method are reasonable:

- Compare the company's blended mortality assumptions against published reinsurer surveys and reinsurance premiums to make sure the assumptions are relatively consistent with the other companies and reinsurers.
- Compare actual experience with the expected experience (the researchers interpret this to be the expected experience resulting from the application of any credibility method), using trend analysis with actuarial judgment.
- Look at the overall mortality assumption shape to make sure it is smooth.
- No formal method for evaluating results of the credibility method, but studies are peer reviewed.
- Mainly based on source of earnings (SOE) analysis. Investigate to understand the root causes of material deviations. Too much weight on industry experience (due to low credibility of internal data) could lead to systematic gains/losses as experience emerges.
- Back testing after obtaining a new assumption. If SOE gain/loss is smaller than before, this indicates that the assumption is more appropriate.
- Calculate A/E based on company experience, then have total company A/E, which can be compared to the final adjustment after normalization. If these are fairly close, the results are considered reasonable.

3.12 Standard Basis for Defining Credibility (Count versus Amount)

Nine out of 10 companies report that credibility factors are calculated on a "by number" basis (as opposed to a "by amount" basis). One of these companies reported that while it uses credibility blending factors by number, it develops assumptions by amount because doing so reflects the financial impact of the assumptions (the researchers interpret this to mean that the weight is developed "by number" but the weights are applied to the assumptions "by amount"). The remaining companies stated that they calculate results both "by number" and "by amount," but chose amount because it captures variability of claim size that better reflects company exposure.

Other comments regarding companies' approaches to defining credibility include:

- Data is broken down as granularly as possible considering the credibility of the experience. Actuarial judgment is mostly applied, specifically in situations where the stakeholder requires more factors.
- Data for different products may be combined to increase credibility where the products are deemed to be sufficiently similar in terms of distribution method, target market, etc.
- For certain products, there is not enough information in industry data regarding all aspects that can impact mortality experience to align the company's experience data with the industry data.
- Internal experience is considered 100% credible for lapse, but views are validated with industry studies when available.
- Using 3,007 claims as criteria for full credibility of mortality, but when amounts differ (interpreted by the researchers to mean that there is significant dispersion in the net amount at risk for each policy in the block under consideration), the number of claims required for full credibility may be much higher. No method established to handle this.
- Using 3,007 claims as a reference for lapse.

3.13 Identification and Treatment of Exceptions

Exceptions involve any instances in which standard credibility methods cannot be applied directly. For example, this may include anomalies in the underlying data, or situations in which the assumptions that underlie a particular method do not hold based on the nature of the data. Nine out of 11 companies surveyed did not describe their approach to identifying and treating exceptions. One company stated that exceptions are identified and treated based on actuarial judgment upon doing impact analysis. One company stated that it looks at anything that could create a difference in the A/E and then adjusts, if needed. There are no exceptions to the method itself, and the approach is consistent across products for lapse and mortality.

Section 4: Analysis and Comparison of Credibility Methods

4.1 General Background and Approach to Analysis

The second component of the research involved performing our own analysis of credibility methods to better understand benefits, limitations and overall results. The initial objective of the credibility analysis was to compare and contrast the approaches/methods that are being utilized in the Canadian marketplace for both mortality and lapse. However, since the survey results made it clear that at this time LFCT is the predominant method, our analysis focused solely on this method.

As described in the survey results above, most of the companies stated they use LFCT to determine credibility for mortality, and some indicated they use LFCT for lapse. Based on the survey responses, no other formal credibility methods are currently being used. Companies that do not use a formal credibility method generally reported using 100% industry data or 100% company data with adjustments. Most of the companies that reported using LFCT indicated they use the "by number" approach (i.e., the Simple Poisson Model) to calculate credibility factors. Several companies reported using 3,007 deaths as the criterion for full credibility, as proposed in the CIA Educational Note. Two companies reported using the "by amount" approach (i.e., the Compound Poisson Model): in one instance to reflect the financial impact of the assumptions, and in the other to capture variability of claim size to better reflect the company exposure. Other industry methods, such as the Bayesian Credibility and the GACT, have been considered by some companies but ultimately rejected because of their complexity. The GACT method was also reported to be avoided because it requires data of several companies and because it does not work well if the random variable has a "heavy" tail (i.e., wide dispersion of the data).

Given these considerations, the researchers worked with the POG to develop an approach to the credibility analysis. The approach ultimately agreed on involved performing an in-depth study of LFCT to understand its benefits and limitations, as well as considerations for companies as they apply this method. We performed analysis:

- With varying parameters (i.e., using a range of expected errors and confidence levels).
- Testing the impacts of using different criteria for defining credibility (i.e., "by count," "by amount").
- Using a range of product types for mortality.
- Using a limited subset of the product types for lapse.
- Using a range of risk classifications for mortality.

These classifications were selected based on judgment, available data and certain limitations related to the LFCT method discussed further below.

The credibility analysis is based on sample data sets developed by the researchers. The development of the sample data sets is described in Section 4.3 below. The spreadsheets used to develop the results by amount are based on the spreadsheet included in the SOA's Credibility Educational Resource for Pension Actuaries, Application of Credibility Theory to Mortality Assumption published in August 2017 ("SOA Pension Credibility Educational Resource").

The researchers did not perform an analysis of the GACT method. However, an overview of the method and comments on considerations regarding the use of GACT are provided in Section 4.2 and Appendix C.

Appendix B includes an annotated bibliography of the documents referenced in this Study. These documents provide good background on credibility theory, and several reference other available resources in their bibliographies. If the document includes a bibliography which references other available resources, this has been noted. The researchers have assigned short names to some of the documents (e.g., CIA Educational Note), which are used to reference the documents throughout the remainder of this Study.

4.2 Overview of Credibility Theory

As noted above, credibility theory may be used to help a company assess whether or not its data is "fully credible" or "100% credible," in which case companies may develop assumptions or create tables based on their own data. If the data is not fully credible, then credibility theory methods may be used to combine the company experience with appropriate base experience (e.g., an industry table or a prescribed valuation table) to develop a more accurate estimate.

Once available company experience data (which may not be fully credible) and appropriate base experience (which is assumed to be fully credible) have been suitably prepared and segmented, they may be blended using credibility weightings.[13]

There are two main approaches to determining the credibility weightings, the GACT and the LFCT. Both approaches strive to produce improved estimates of future events based on combining company experience data and appropriate base experience. Both approaches use the following linear estimator formula to combine the company experience and the base experience:[14]

$$X_E = Z\overline{X} + (1-Z)\mu \qquad \text{(Formula 1)}$$

where

- X_E is estimated based on the combined experience.
- Z is the credibility factor or weighting given to the sample data (i.e., the company experience data).
- $\overline{X}$ is the mean calculated from the company experience data.

[13] CIA Educational Note, Section 500.
[14] CIA Educational Note, Sections 530 and 540; SOA Pension Credibility Educational Resource Section 2.2.

- μ is the mean of the underlying distribution (i.e., the "population mean," which is assumed to be the base experience).

If the company experience is deemed to be fully credible, Z is set to 1.0. The CIA Educational Note states "Full credibility means it is appropriate to use only the portfolio's own experience and to ignore the entire industry data."[15] If Z is equal to 0, then no weight is assigned to the company's experience. If Z is between 0 and 1, then the formula provides the method for weighting the two sets of experience data.

The difference between the two methods is how *Z* is determined. We reviewed other available sources of information comparing the two methods.

The SOA Credibility Theory Practices Report provides this summary of the differences:

> In both the Limited Fluctuation and the Bühlmann Empirical Bayesian methods, the results are calculated with respect to a mean (A/E ratio) and incorporate a variance. The methods differ in the treatment of the components of the variance (σ^2). The total variance of the observations is the sum over all companies of two different sources of variation, which are:
>
> 1. For each company, the variation of a company's observations about that company's mean, and
> 2. The variation between each company's mean and the overall mean.
>
> Limited Fluctuation credibility uses only the first source, while the Bühlmann Empirical Bayesian method uses both. Thus, Limited Fluctuation credibility only requires data from the company being studied. For the Bühlmann Empirical Bayesian approach, data is required for all companies under study.[16]

The CIA Educational Note provides general information about the development of the LFCT and the GACT credibility methods. It also includes appendices with additional detailed information regarding the development of the formulas used in these methods. The SOA Pension Credibility Educational Resource provides an overview of GACT and detailed information about the development of LFCT. It states that LFCT has a weaker theoretical basis and requires subjective choices, but it is more practical to apply. GACT has stronger theoretical support but requires information that may not be available or not worth the collection effort.[17] The SOA Credibility Theory Practices Report demonstrates the development of the LFCT formulas and the GACT formulas using A/E ratios. Chapter 8 of the Casualty Actuarial Society (CAS) Textbook referenced in Appendix B ("CAS Textbook") has been used by the SOA for its preliminary exam covering credibility, and provides additional background on these methods. The CIA Standards of Practice include points which Canadian actuaries must demonstrate they have considered in assessing credibility.[18] Actuarial Standard of Practice 25: Credibility Procedures includes professional standards related to credibility procedures for members of the AAA, and the AAA Credibility Practice Note provides information to actuaries on current and emerging practices related to credibility procedures. Given that these other sources of background information for credibility methods are readily available, the researchers have not reproduced the information in this paper. Instead, the researchers have limited the information included in this paper to that deemed necessary for the reader to obtain an understanding of the methods and their uses and limitations. The source of the information included in the descriptions of the methods that follow is noted in the text itself, or the footnotes, or both. Some information varies slightly from the noted references due to the researchers' use of the terms "company experience data" and "base experience" to describe the sample data and the population data respectively. The researchers have relied heavily on the CIA Educational Note since all companies responding to the survey indicated that they use this as their guideline for applying credibility theory. The CIA Educational Note is specific to developing expected mortality for individual life insurance business. However, many of the concepts covered in the Educational Note are useful in the broader application of credibility theory.

Further details regarding the two primary methods can be found in Appendix C.

---

[15] CIA Educational Note, Section 540, page 15.
[16] SOA Credibility Theory Practices Report, page I.21.
[17] SOA Pension Credibility Educational Resource, Section 2.2.
[18] CIA Standards of Practice, Section 1620.

4.3 Development of the Sample Data Sets

In order to perform credibility analysis using LFCT, the researchers developed representative sample data sets, which refer to hypothetical historical experience data constructed by the researchers intended to demonstrate representative industry mortality and lapse experience.

Sample Data Sets for Mortality

The sample data sets for mortality were developed using data from the CIA Mortality Study, "Canadian Standard Ordinary Life Experience 2014–2015, Using 86–92 Tables,"[19] developed by the Research Executive Committee's Experience Studies Subcommittee and published in July 2017 (Document 217077, hereinafter referred to as the "CIA Mortality Study").

Data sets were developed for the product types listed below. For each product type, two data sets were developed: one data set includes all of the underlying data available in the mortality study ("Full Data");[20] the other data set aggregates the data by attained age ("AA Totals"):

- Whole Life – with face size less than $100K (WL with <$100k).[21]
- Whole Life – with face size of $100K+ (WL with >=$100k).[22]
- Renewable Term with 10-year renewal term (T10).
- Renewable Term with 20-year renewal term (T20).
- Universal Life with YRT Cost of Insurance (UL YRT).
- Universal Life with Level Cost of Insurance (UL LCOI).
- Term-to-100.

Data sets were also developed for the following risk classifications using all of the underlying data available in the mortality study ("Full Data"):

- Male/Female.
- Smoker/Non-smoker.
- Standard/Preferred (includes preferred standard and super preferred).

Sample Data Sets for Lapse

The researchers had concerns with applying the LFCT method to lapse for reasons that are discussed in Section 4.5. Thus, a more limited analysis was performed using lapse data.

Data sets were developed for the following product types for lapse:

- Whole Life – with face size of $100K+.
- Term-to-100.

In developing these data sets, the researchers used the exposure data from the CIA Mortality Study and applied lapse rates from other industry studies.

[19] The researchers note that 100% of the CIA 86–92 tables is used as the expected mortality for the analysis. It is our understanding that the mortality assumption for most companies may be based on the CIA 97–04 tables. The researchers anticipate that use of the 97-04 tables as the expected basis may change the number of deaths needed for full credibility "by amount." However, the researchers would not expect this to impact the overall observations from the study. The analysis spreadsheets include instructions for calculating results for a different expected basis. However, it is important to note that the spreadsheets are intended to illustrate the calculation of the number of deaths needed for full credibility "by amount" for the LFCT method for a sample data set and to illustrate how the results might vary by assumption, product type and risk classification. Actual results for a company will depend on the nature of the company's block of business.

[20] It is important to note that the researchers did not have access to the seriatim data. The "Full Data" reflects the most granular data from the CIA Mortality Study.

[21] $ denotes Canadian dollars.

[22] $ denotes Canadian dollars.

- For the whole life data set, the researchers used the lapse rates in the CIA Research Paper "Lapse Experience Under Universal Life Level Cost of Insurance Policies" developed by the Research Committee's Individual Life Experience Subcommittee and published in September 2015 (Document 215076, hereinafter referred to as the "CIA UL LCOI Lapse Study"). This study was used because the researchers were unable to find a CIA lapse study for whole life. The researchers applied the lapse rates in the CIA UL LCOI Lapse Study to the exposure from the CIA Mortality Study included in the Whole Life – with face size of $100K+ mortality data set.
- For the Term-to-100 data set, the researchers found a recent CIA Term-to-100 lapse study, the CIA Research Paper "Lapse Experience Under Term-to-100 Insurance Policies" developed by the Research Committee's Individual Life Experience Subcommittee and published in September 2015 (Document 215075, hereinafter referred to as the "CIA Term-to-100 Lapse Study"). The researchers used the lapses from this study applied to the exposures for the Term-to-100 product in the CIA Mortality Study to develop the data set. The exposures from the CIA Term-to-100 Lapse Study could have been used for the analysis, but the researchers opted to use the CIA Mortality Study exposures for consistency.

For each product type, two sample data sets were developed: one which includes all of the underlying data available in the mortality study ("Full Data"), and one which aggregates the data by duration ("D Totals").

4.4 Credibility Analysis for Mortality

The credibility analysis for mortality involves a comparison of the number of deaths needed for full credibility when defining credibility "by count" and "by amount" for the LFCT method using the sample data sets developed for mortality described in Section 4.3.[23]

As noted in the overview of the LFCT method in Section 4.2, variations in claim size are ignored in the Simple Poisson Model. The researchers refer to this criterion for defining credibility as "by policy" or "by count." The Compound Poisson Model incorporates the effect of variation in claim size. The researchers refer to this criterion for defining credibility as "by amount."

For purposes of the credibility analysis for mortality, the researchers have assumed that it is reasonable to approximate the Poisson distribution with a Binomial distribution at all attained ages, including higher attained ages. See the subsection titled "Approximating the Binomial with a Poisson" in Section 4.5 below for additional discussion regarding this assumption.

Analysis Spreadsheets

The credibility analysis for mortality includes a range of confidence levels/margins of error, product types and risk classifications. The analysis for each product type and risk classification is included in the following spreadsheets:

Product Type

- Analysis – Whole Life less than $100K.
- Analysis – Whole Life $100K+.
- Analysis – 10-year renewable term.
- Analysis – 20-year renewable term.
- Analysis – UL YRT.
- Analysis – UL LCOI.
- Analysis – Term-to-100.

Risk Classification

- Analysis – Male.
- Analysis – Female.

[23] In conducting the analysis for mortality using the CIA Mortality Study, the expected mortality is based on the CIA 86–92 tables.

- Analysis – Smoker.
- Analysis – Non-smoker.
- Analysis – Standard.
- Analysis – Preferred.

For the different product types, the number of deaths needed for full credibility are summarized for each combination of confidence level/margin of error for the “Full Data” and the “AA Totals” on separate tabs labeled “Mortality Summary-Full Data” and “Mortality Summary-AA Totals” in the analysis spreadsheets. For the different risk classifications, the results are only based on the “Full Data” so there is only one summary tab. The results “by amount” shown in the summary tabs may be reproduced by accessing the spreadsheet (“calculator”) in the “Calc” tab adjacent to the summary tab and entering the desired margin of error and confidence level. Given that the Simple Poisson Model ignores variations in claim size, the number of deaths needed for full credibility “by count” for a particular confidence level/margin of error does not vary by product type or risk classification, and the results are the same as those given in the table labeled “Standard Normal Table – Range and Probability Parameters” from the CIA Educational Note reprinted in Section 4.2. Even so, the results “by count” are calculated in the summary tabs. The summary tabs also include the number of deaths for various levels of partial credibility and summary statistics that may be helpful in comparing results.[24]

The spreadsheets used to develop the results “by amount” are based on the spreadsheet included in the SOA Pension Credibility Educational Resource. The detailed development of the formula used in the spreadsheet for calculating the number of deaths needed for full credibility “by amount” is shown in the Appendix of the SOA Pension Credibility Educational Resource. Although different symbols are used, this formula is identical to the formula for the Compound Poisson Model on page 40 of the CIA Educational Note, except that the formula in the CIA Educational Note assumes r=3% and p=90%, whereas the formula in the SOA Pension Credibility Educational Resource is generalized. The sample data used to calculate the results “by amount” is included in separate tabs within the analysis spreadsheet. It is the intent of the researchers that readers be able to reproduce the results in the summary tabs from the sample data.

The SOA Pension Credibility Educational Resource discusses the construction of mortality tables and makes the point that to build a mortality table from scratch, fully credible data would be needed at each age. It further states that a more practical approach would be to take an existing standard mortality table and adjust it using the LFCT methodology.[25] The SOA Pension Credibility Educational Resource explains that

> [t]he LFCT adjustment works to ‘shift’ the standard mortality table up or down based on the plan’s experience. The overarching assumption for this purpose is that the true mortality table for the subject plan is a constant multiple of the standard table. It is assumed the same multiplier is applied at all ages, so the shape of the new table is the same as the underlying standard table. This is why, when selecting the standard table to use in an experience study, it is important to consider the shape of the standard table compared to the shape of the actual experience for the plan being valued.[26]

The SOA Pension Credibility Educational Resource also develops the estimator of the multiple that will shift the entire mortality table, and the functionality for calculating the multiple is included in the spreadsheet. Although this multiple is not the subject of the analysis, the researchers have also included this functionality in the analysis spreadsheets noted above.

Analysis Results

The credibility analysis for mortality involves a comparison of the number of deaths needed for full credibility when defining credibility “by count” (i.e., the Simple Poisson Model) and “by amount” (i.e., the Compound Poisson Model) for the LFCT method using a range of confidence levels (labeled “p”) and margins of error (labeled “r”) for various product types and risk classifications.

[24] The results by count in the analysis spreadsheets for partial credibility differ slightly from those noted in the CIA Educational Note due to rounding.
[25] SOA Pension Credibility Educational Resource, Section 3.2.
[26] SOA Pension Credibility Educational Resource, Section 3.4.

Product Type

Below is a summary of the results of the analysis by product type for the "Full Data."

Number of Deaths Needed for Full Credibility based on selected values of r and p – By Product									
Analysis done on:		By Count*	By Amount**						
	p	All products	WL with <$100K	WL with >=$100K	T10	T20	UL YRT	UL LCOI	T100
r = 1%	90%	27,060	61,382	105,856	76,097	63,419	244,212	186,254	220,626
	95%	38,416	87,153	150,299	108,046	90,046	346,743	264,453	313,255
	99%	66,358	150,529	259,593	186,616	155,525	598,888	456,758	541,048
r = 3%	90%	**3,007**	6,820	11,762	8,455	7,047	27,135	20,695	24,514
	95%	4,268	9,684	16,700	12,005	10,005	38,527	29,384	34,806
	99%	7,373	16,725	28,844	20,735	17,281	66,543	50,751	60,166
r = 5%	90%	1,082	2,455	4,234	3,044	2,537	9,768	7,450	8,825
	95%	1,537	3,486	6,012	4,322	3,602	13,870	10,578	12,530
	99%	2,654	6,021	10,384	7,465	6,221	23,956	18,270	21,642

* The "By Count" results reflect the number of deaths needed for full credibility for the noted confidence levels (p) and margins of error (r), using the Simple Poisson Model. A minimum of 3,007 deaths is recommended for full credibility in the CIA Educational Note.[27]

** The "By Amount" results reflect the number of deaths needed for full credibility for the noted confidence levels (p) and margins of error (r) using the Compound Poisson Model.

[27] CIA Educational Note, page 16, item 3.

The researchers offer the following observations:

1. The number of deaths needed for full credibility "by amount" are materially higher than the number of deaths needed for full credibility "by count" for all product types for the sample mortality data sets. This result is not unexpected given that the underlying data for the sample data sets is the CIA mortality data and as such there is significant variation in the exposure (in other words, there is a wide range of face amounts across companies included in the CIA Mortality Study) in the blocks under consideration.[28]
2. As expected, the number of deaths needed for full credibility "by count" does not vary by product type. This is due to the nature of the Simple Poisson Model which is used to develop the results for the LFCT method "by count." The only parameters which impact the number of occurrences (in this case deaths) needed for full credibility are the confidence level and margin of error selected. This result is discussed further in Appendix C and in the CIA Educational Note.[29]
3. The number of deaths needed for full credibility "by count" and "by amount" both vary depending on the confidence level and the margin of error. Decreasing the margin of error results in a much larger increase in the number of deaths needed for full credibility than increasing the confidence level both "by count" and "by amount" for all product types for the sample data sets.
4. The number of deaths needed for full credibility "by amount" varies by product type for the sample mortality data sets (e.g., T10 vs T100 vs UL YRT). These results are driven in part by the relative variation in the exposure for each of these data sets (in other words, products with greater variation in face amount, all else equal, will require more deaths by amount than products with less variation). The number of deaths needed for full credibility varies depending on the characteristics of the underlying portfolio and the assumptions used.
5. The researchers have not reproduced excerpts of the results of the analysis by product for the data aggregated by attained age "AA Totals." However, the researchers note that for all product types the number of deaths needed for full credibility "by amount" is materially less when the exposures are aggregated by attained age. This simplified approach masks some of the variability in the underlying data because it assumes that the exposure and expected mortality are the same for all lives at a particular attained age.
6. These results are based on the sample data sets constructed by the researchers using all of the industry mortality data available. Actual results for a company will depend on the nature of the company's block of business, including the amount of variation in net amount at risk for the company's block.

[28] CIA Educational Note, page 16, item 2.
[29] CIA Educational Note, page 16, items 2 and 3.

Risk Classification

Below is a summary of the results of the analysis by risk classification for the "Full Data."

Number of Deaths Needed for Full Credibility based on selected values of r and p – By Risk									
Analysis done on:		By Count*	By Amount**						
	P	All risk classifications	Male	Female	Smoker	Non-Smoker	Standard	Preferred
r = 1%	90%	27,060	280,454	373,421	187,507	174,617	354,788	65,985
	95%	38,416	398,202	530,201	266,231	247,930	503,745	93,689
	99%	66,358	687,767	915,753	459,829	428,220	870,059	161,871
r = 3%	90%	**3,007**	31,162	41,491	20,834	19,402	39,421	7,332
	95%	4,268	44,245	58,911	29,851	27,548	55,972	10,410
	99%	7,373	76,419	101,750	51,092	47,580	96,673	17,980
r = 5%	90%	1,082	11,218	14,937	7,500	6,985	14,192	2,693
	95%	1,537	15,928	21,208	10,649	9,917	20,150	3,748
	99%	2,654	27,511	36,630	18,393	17,129	34,802	6,473

* The "By Count" results reflect the number of deaths needed for full credibility for the noted confidence levels (p) and margins of error (r), using the Simple Poisson Model. A minimum of 3,007 deaths is recommended for full credibility in the CIA Educational Note[30]

** The "By Amount" results reflect the number of deaths needed for full credibility for the noted confidence levels (p) and margins of error (r) using the Compound Poisson Model.

The researchers offer the following observations (many of which are similar to the results by product type above):

1. The number of deaths needed for full credibility "by amount" are materially higher than the number of deaths needed for full credibility "by count" for all risk classifications for the sample mortality data sets. This result is not unexpected given that the underlying data for the sample data sets is the CIA mortality data and as such there is significant dispersion in the exposure for each policy in the block under consideration.[31]
2. As expected, the number of deaths needed for full credibility "by count" does not vary by risk classification. This is due to the nature of the Simple Poisson Model which is used to develop the results for the LFCT method "by count." The only parameters which impact the number of occurrences (in this case deaths) needed for full credibility are the confidence level and margin of error selected. This result is discussed further in Appendix C and in the CIA Educational Note.[32]
3. The number of deaths needed for full credibility "by count" and "by amount" both vary depending on the confidence level and the margin of error. Decreasing the margin of error results in a much larger increase in the number of deaths needed for full credibility than increasing the confidence level both "by count" and "by amount" for all risk classifications for the sample data sets.

[30] CIA Educational Note, page 16, item 3.
[31] CIA Educational Note, page 16, item 2.
[32] CIA Educational Note, page 16, items 2 and 3.

4. The number of deaths needed for full credibility "by amount" varies by risk classification for the sample mortality data sets, with the results for females being somewhat higher than for males, the results for smokers being slightly higher than for non-smokers and the results for the standard class being materially higher than for the preferred class. These results are driven in part by the relative variation in the exposure for each of these data sets. The number of deaths needed for full credibility varies depending on the characteristics of the underlying portfolio and the assumptions used.
5. These results are based on the sample data sets constructed by the researchers. Actual results for a company will depend on the nature of the company's block of business, including the amount of variation in net amount at risk for the company's block.

4.5 Credibility Analysis for Lapse

The credibility analysis for lapse involves a comparison of the number of lapses needed for full credibility when defining credibility "by count" and "by amount" for the LFCT method using the sample data sets developed for lapse described in Section 4.3.

The researchers had concerns with applying LFCT to lapse given the need to approximate a Binomial distribution with the Poisson distribution for the formulas being used for the analysis (which are consistent with the formulas in the CIA Educational Note). See below for further explanation. However, several survey respondents indicated that they use LFCT for lapse and one respondent found this method to be the most representative of the company's experience and view on tail risks. This company stated that the method allows it to recognize more rapidly its recent experience as experience emerges. Thus, the POG encouraged the researchers to pursue this analysis, at least for some product types where the lapse rates tend to be lower. However, the researchers note, and caution readers, that this is an evolving area of practice.

Approximating the Binomial with a Poisson

As noted in Section 4.2, the CIA Educational Note makes the following point:[33]

> Although the theoretical distribution for mortality is binomial, when the probabilities of the event (death, represented by the random variable X in the above formulas) are small, the Poisson distribution provides a reasonable approximation to a binomial distribution.

The question of whether or not the probabilities of the event are small enough in order for the Poisson to provide a reasonable approximation to a Binomial distribution is an important consideration and potential limitation of this approach. The researchers believe that this is a particularly important item to consider for the lapse assumption.

How small must *p* be in order for the Poisson distribution to be a good approximation to a Binomial distribution?

Recall that for a Binomial distribution the mean, μ, equals np and the variance, σ^2, equals npq, where p is the probability of "success" from trial to trial for the Binomial distribution and q is 1-p. For a Poisson distribution, the mean, μ, equals np and the variance, σ^2, equals np. In each case, n is the "number of trials" and p is the "probability of success" (i.e., the probability of the event occurring) from trial to trial. The approximation assumes that p is small enough such that it is reasonable to assume that q=1-p is approximately equal to 1 or that variances of the two distributions are approximately equal (np is approximately equal to npq). In other words, it must be reasonable to assume that p is approximately equal to 0. This assumption may not be reasonably met for higher attained ages for mortality and in certain durations (e.g., early durations or term renewal) for lapse.

For those situations in which *n* is large and *p* is very small, the Poisson distribution can be used to approximate the Binomial distribution. The larger the *n* and the smaller the *p*, the better is the approximation. There are several common rules of thumb for assessing whether or not it is reasonable to assume that the Poisson is a good approximation to a Binomial:[34]

[33] CIA Educational Note, page 16, item 1.
[34] NIST/SEMATECH, "6.3.3.1. Counts Control Charts", *e-Handbook of Statistical Methods.*

- N >= 20 and p < .05
- N >= 100 and p <= .01

The researchers reviewed industry lapse rates for each of the product types included in the mortality analysis. The lapse rates did not meet the rules of thumb noted above at all durations for any product type. However, the lapse rates for Whole Life 100k+ and Term-to-100 appear to be reasonably "small" for many durations. So, the credibility analysis was performed for these two product types.

An Excel spreadsheet may easily be used to calculate the probabilities for the Binomial distribution and the Poisson distribution for a given level of n and p and to compare these differences in the probabilities. The researchers believe that it will ultimately require actuarial judgment to decide if the result is such that it is reasonable to use the Poisson distribution to approximate the Binomial distribution.

The researchers note that if the assumption is not met, then formulas may be derived without using the simplifying assumption, which is the approach taken in the SOA Credibility Theory Practices Report.[35] Alternatively, an adjustment may be made to the standard for full credibility (as derived in the formulas below).

Chapter 8 of the CAS Textbook includes the following general formula for determining the standard for full credibility for frequency (i.e., "by count") when the Poisson assumption does not apply:[36]

> 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 $\sigma_f^2/\mu_f = 1$.

Chapter 8 of the CAS Textbook includes the following general formula for determining the standard for full credibility for pure premiums (i.e., "by amount") when the Poisson assumption does not apply:[37]

> *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]
>
> $$n_F = \{y^2/k^2\}(\sigma_f^2/\mu_f + \sigma_s^2/\mu_s^2), \qquad (2.5.5)$$
>
> 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).

where

y = z-score associated with the desired confidence level (i.e., p).

---

[35] SOA Credibility Theory Practices Report, page I.6.
[36] Chapter 8 of the CAS Textbook.
[37] Chapter 8 of the CAS Textbook.

k = the desired margin of error (i.e., r).

f denotes frequency.

s denotes severity.

n_F denotes the number of claims needed for full credibility.

If the variance of the distribution is larger than the mean, then the standard for full credibility is higher than it is in the Poisson case. If the variance of the distribution is less than the mean, then the standard for full credibility is lower than it is in the Poisson case.

Given that the variance of the Binomial is smaller than the mean, the standard for full credibility is lower than it is in the Poisson case (i.e., the standard for full credibility will be higher if the Poisson is used to approximate the Binomial).

Analysis Spreadsheets

The credibility analysis for lapse includes a range of confidence levels/margins of error and is focused on two product types where the researchers deemed it was not unreasonable to use the Poisson distribution to approximate the Binomial distribution. The analysis for each product type is included in the following spreadsheets:

Product Type

- Analysis – Whole Life $100K+.
- Analysis – Term-to-100.

For the different product types, the number of lapses needed for full credibility are summarized for each combination of level of confidence/margin of error for the "Full Data" and the duration totals "D Totals" on separate tabs labeled "Lapse Summary-Full Data" and "Lapse Summary-D Totals" in the analysis spreadsheets. The results "by amount" shown in the summary tabs may be reproduced by accessing the spreadsheet ("calculator") in the "Calc" tab adjacent to the summary tab and entering the desired margin of error and confidence level. Given that the Simple Poisson Model ignores variations in claim size, the number of deaths needed for full credibility "by count" for a particular confidence level/margin of error does not vary by product type and the results are the same as those given in the table labeled "Standard Normal Table – Range and Probability Parameters" from the CIA Educational Note reprinted in Section 4.2. Even so, the results "by count" are calculated in the summary tabs. The summary tabs also include the number of lapses for various levels of partial credibility and sample statistics that may be helpful in comparing results.

As with the mortality credibility analysis, the spreadsheets used to develop the results "by amount" are based on the spreadsheet included in the SOA Pension Credibility Educational Resource.

Analysis Results

The results of the credibility analysis for lapse display the number of lapses needed for full credibility when defining credibility "by count" (i.e., the Simple Poisson Model) and "by amount" (the Compound Poisson Model) for the LFCT method using a range of confidence levels (labeled "p") and margins of error (labeled "r") for two product types where the researchers deemed it was not unreasonable to use the Poisson distribution to approximate the Binomial distribution.

Product Type

Below is a summary of the results of the analysis for lapse by product type for the "Full Data."

Number of Lapses Needed for Full Credibility based on selected values of r and p				
Analysis done on:		By Count*	By Amount**	
	p	All products	WL with >$100K	T100
r = 1%	90%	27,060	148,813	172,689
	95%	38,416	211,292	245,192
	99%	66,358	364,940	423,491
r = 3%	90%	3,007	16,535	19,188
	95%	4,268	23,477	27,244
	99%	7,373	40,549	47,055
r = 5%	90%	1,082	5,953	6,908
	95%	1,537	8,452	9,808
	99%	2,654	14,598	16,940

The researchers have also included the summaries for mortality analysis (from Section 4.4 above) of the Whole Life $100k+ and Term-to-100 below for comparison:

Number of Deaths Needed for Full Credibility based on selected values of r and p				
Analysis done on:		By Count*	By Amount**	
	p	All products	WL with >=$100K	T100
r = 1%	90%	27,060	105,856	220,626
	95%	38,416	150,299	313,255
	99%	66,358	259,593	541,048
r = 3%	90%	**3,007**	11,762	24,514
	95%	4,268	16,700	34,806
	99%	7,373	28,844	60,166
r = 5%	90%	1,082	4,234	8,825
	95%	1,537	6,012	12,530
	99%	2,654	10,384	21,642

* The "By Count" results reflect the number of lapses/deaths needed for full credibility for the noted confidence levels (p) and margins of error (r), using the Simple Poisson Model. A minimum of 3,007 deaths is recommended for full credibility in the CIA Educational Note.[38]

** The "By Amount" results reflect the number of lapses/deaths needed for full credibility for the noted confidence levels (p) and margins of error (r) using the Compound Poisson Model.

The researchers offer the following observations (many of which are similar to the results for mortality above):

1. The number of lapses needed for full credibility "by amount" are materially higher than the number of lapses needed for full credibility "by count" for both product types for the sample lapse data sets. This result is not unexpected given that the underlying exposure data for the sample data sets is the CIA mortality data and as such there is significant dispersion in the exposure for each policy in the block under consideration.[39]
2. As expected, the number of lapses needed for full credibility "by count" do not vary by product type. This is due to the nature of the Simple Poisson Model which is used to develop the results for the LFCT "by count." The only parameters which impact the number of occurrences (in this case lapses) needed for full credibility are the confidence level and margin of error selected. This result is discussed further in Appendix C and in the CIA Educational Note.[40]
3. The number of lapses needed for full credibility "by count" and "by amount" both vary depending on the confidence level and the margin of error. Decreasing the margin of error results in a much larger increase in the number of lapses needed for full credibility than increasing the confidence level both "by count" and "by amount" for both product types for the sample data sets.

[38] CIA Educational Note, page 16, item 3.
[39] CIA Educational Note, page 16, item 2.
[40] CIA Educational Note, page 16, items 2 and 3.

4. The number of lapses needed for full credibility “by amount” varies by product type for the sample lapse data sets, with the results for Term-to-100 being modestly higher than for Whole Life $100k+. These results are driven in part by the relative variation in the exposure for each of these data sets. The number of lapses needed for full credibility varies depending on the characteristics of the underlying portfolio and the assumptions used.
5. The number of lapses needed for full credibility is higher for whole life than the number of deaths needed for full credibility for whole life for the sample data sets. This result is driven at least in part by the difference in the expected mortality rates and lapse rates for the sample data sets. The data sets were constructed so that the underlying exposures would be the same for comparison purposes.
6. The number of lapses needed for full credibility is lower for Term-to-100 than the number of deaths needed for full credibility for Term-to-100 for the sample data sets. This result is driven at least in part by the difference in the expected mortality rates and lapse rates for the sample data sets. The data sets were constructed so that the underlying exposures would be the same for comparison purposes.
7. The researchers have not reproduced excerpts of the results of the analysis by product for the data aggregated by duration “D Totals.” However, the researchers note that for both product types the number of lapses needed for full credibility “by amount” is materially less when the exposures are aggregated by duration. This simplified approach masks some of the variability in the underlying data because it assumes that the exposure is the same for all lives at a particular duration.
8. These results are based on the sample data sets constructed by the researchers. Actual results for a company will depend on the nature of the company’s block of business, including the amount of variation in net amount at risk for the company’s block.

Section 5: Acknowledgements

The authors' deepest gratitude goes out to those without whose efforts this project could not have come to fruition:

- Without the companies that responded to the survey, whose names are included below, this study would not have been possible.

  BMO Life Assurance Company

  Desjardins Financial Security Life Assurance Company

  The Empire Life Insurance Company

  Equitable Life Insurance Company of Canada

  Foresters Life Insurance and Annuity Company

  Industrial Alliance (iA Financial Group)

  Ivari

  La Capitale Insurance and Financial Services

  Manulife

  RBC Insurance

  Sun Life Assurance Company of Canada

- The POG and SOA/CIA representatives for their diligent work overseeing questionnaire development, analyzing respondent answers, assessing the analysis performed and reviewing and editing this report for accuracy and relevance.

  Project Oversight Group Members

  Vera Ljucovic FSA, FCIA (Chair)

  Steven Ekblad FSA, MAAA*

  Annie Girard, FSA, FCIA*

  Frédérick Guillot*

  Carole Vincent FSA, FCIA*

  Yu Luo ASA, MAAA, PhD*

  SOA representatives

  Leonard Mangini FSA, MAAA

  Ben Marshall FSA, FCIA, CERA, MAAA, J.D.

  Jan Schuh

  Ronora Stryker ASA, MAAA

  CIA Experience Research Committee Liaison

  Carrie Lam, FSA, CERA, FCIA

Appendix A: Survey Questions

1. How is your Company applying credibility theory for mortality and lapse assumption setting?
   a) What are the primary credibility methods used? If the methods are different for mortality and lapse, please indicate. If you make any Company-specific adjustments to standard industry methods, please describe those.
   b) How is the credibility method typically selected? What are the key drivers of the choice of methodology? Include items such as the nature of the experience data, the application (pricing, reserving, etc.), data and/or resource constraints, regulatory constraints, and system or model limitations.
   c) What are the benefits and drawbacks of the methods you are using? If you are able, please also comment on benefits and drawbacks of other industry methods, even if you do not use them.
   d) What issues have you encountered with the methods and what approaches do you take to dealing with these issues?
   e) How does your approach to selecting the credibility method balance the need for sufficient data (i.e. longer time periods) and recent data (i.e. shorter time periods)?
   f) Does the credibility method vary by use? For example, does the method vary depending on whether the assumption is being set for pricing, reserving or overall company risk assessment?
2. Is your Company applying credibility theory to determine other assumptions? If so, please provide an overview of the types of assumptions to which credibility theory is applied.
   a) Are there unique aspects for these assumptions that make the process different than what is described in question 1 for mortality and lapse? If so, please describe.
3. What internal and external guidelines (regulatory, professional, company-specific, etc.) does your Company use when applying credibility theory?
4. What data does your Company use when applying credibility theory? How does the data vary by use? How is the data aggregated? For example, are there differences in how data is collected, aggregated, and analyzed depending on which assumption is under evaluation?
5. What software package(s) does your Company use to analyze the credibility of data, and for what purpose(s) (e.g. data organization, actual-to-expected analysis, application of a credibility methodology)?
   a) Why was this software chosen?
   b) What do you consider to be the benefits and drawbacks of the software chosen, compared to other options available?
6. How does the application of credibility theory vary by product type and any other factors (for example underwriting method, size of block, etc.)?
7. How are adjustments made for new products or changes in underwriting criteria, where directly relevant data is not available?
8. If your Company supplements experience data with industry/population data, please describe instances in which this occurs.
   a) How does your Company adjust for basis risk (i.e. the risk that the differences in populations may result in an inappropriate assumption)?
   b) How do you weight Company experience data that is not fully credible with industry or population data?
9. What types of analyses does your Company perform to evaluate whether the results of a credibility method are reasonable?
10. Does your Company have standard criteria that define credibility (e.g. based on number of deaths or face amount for mortality, based on the item being evaluated)? If so, please describe the criteria and how they were set. (e.g. For mortality, are they set based on the Canadian Institute of Actuaries Educational Note Document 202037, *Expected Mortality: Fully Underwritten Canadian Individual Life Insurance Policies* (attached for ease of reference)? For lapse, are there internal guidelines, etc.?) Although not intended to be all-inclusive, below are some examples of questions you may wish to consider in your response:

a) Are credibility factors calculated on a 'by number'; or 'by amount' basis, and why was that basis chosen?
b) How are other factors, such as size of data set, data distribution (volatility/dispersion of underlying data), product mix, geographic mix, distribution channel mix, etc. considered, if at all?
c) For mortality, if using the 'by amount' basis, considering that the referenced CIA Educational Note does not give specific guidance other than what is included in item 5 on page 17,[41] specifically describe how your Company assigns credibility using this basis.

11. If your Company has standard criteria that define credibility, how are exceptions identified and treated (for example, due to significant dispersion in the net amount at risk)?
12. Please share any other relevant information regarding credibility application at your Company that may not have been already requested.

[41] CIA Educational Note, item 5, page 17: "The parameters defined in Step 4 above are suggested for use in most situations. A significant dispersion in net amount at risk in the inforce block will increase volatility and could result in the need to use a higher number of deaths."

Appendix B: Annotated Credibility Bibliography

This is a list of the documents referenced in this Study. These documents provide good background on credibility theory. Several reference other available resources in their bibliographies; if so, this has been noted.

1. CIA Educational Note: Canadian Institute of Actuaries Committee on Life Insurance Financial Reporting. "Expected Mortality: Fully Underwritten Canadian Individual Life Insurance Policies." CIA Educational Note (Document 202037), July 2002, www.actuaries.ca/members/publications/2002/202037e.pdf. (Section 500 includes sources of information. Appendices 1–3 include additional information related to probability and statistical concepts, Limited Fluctuation Credibility Theory, and Greatest Accuracy Credibility Theory/Bühlmann Method respectively.)
2. SOA Pension Credibility Educational Resource: Irina Pogrebivsky. "Credibility Educational Resource for Pension Actuaries, Application of Credibility Theory to Mortality Assumption." Research Paper, Society of Actuaries, 2017, www.soa.org/files/static-pages/sections/pension/credibility-resource-pension.pdf. (Appendix 5.1 contains a bibliography of other resources.)
3. SOA Credibility Theory Practices Report: Klugman, Stuart, Tom Rhodes, Marianne Purushotham and Stacy Gill. "SOA Credibility Theory Practices Report." Research Paper, Society of Actuaries, 2009, www.soa.org/research-reports/2009/research-credibility-theory-pract/. (Section III of this paper contains a bibliography of other resources.)
4. AAA Life Valuation Subcommittee. "Credibility Practice Note." Public Policy Practice Note, revised July 2008, http://actuary.org/files/publications/Practice_note_on_applying_credibility_theory_july2008.pdf. (Appendix 5.1 has an extensive bibliography of additional credibility resources.)
5. Credibility Task Force of the General Committee of the Actuarial Standards Board. "Actuarial Standard of Practice 25: Credibility Procedures." Revised December 2013, www.actuarialstandardsboard.org/wp-content/uploads/2014/02/asop025_174.pdf. (Appendix 1 provides a high-level background discussion on credibility practice.)
6. Chapter 8 of the CAS Textbook: Mahler, Howard C., and Curtis Gary Dean. "Credibility." Chapter 8 in *Foundations of Casualty Actuarial Science*. Casualty Actuarial Society, 2001, www.soa.org/files/pdf/C-21-01.pdf
7. SOA *Financial Reporter* PBA Corner Credibility Article: Karen Rudolph, Ruijuan Wang. "PBA Corner." Article from *The Financial Reporter*, June 2016, Issue 105, www.soa.org/Library/Newsletters/Financial-Reporter/2016/june/fr-2016-iss105-rudolph-wang.aspx
8. National Association of Insurance Commissioners, Valuation Manual, 2019 Edition, www.naic.org/documents/cmte_a_latf_related_val_2019_edition.pdf
9. AAA Group Long-Term Disability Valuation Standard: American Academy of Actuaries, "Group Long-Term Disability Valuation Standard Report of the American Academy of Actuaries' Group Long-Term Disability Work Group." Presented to the National Association of Insurance Commissioners' Health Actuarial Task Force, October 2013, www.actuary.org/files/Final_GLTDWG_Table_Report_Final_Version_Oct3_0.pdf
10. CIA Standards of Practice: Canadian Institute of Actuaries, Standards of Practice, Section 1620. Effective March 1, 2019, www.cia-ica.ca/docs/default-source/standards/sc030119e.pdf

Appendix C: Credibility Methods

Greatest Accuracy Credibility Theory

The SOA Pension Credibility Educational Resource makes the following observations about GACT:

> This method attempts to produce estimates that minimize the expected value of the square of the difference between the estimate and the quantity being estimated. In this way it endeavors to optimize the weights so credibility is determined based on both the 'accuracy' of relevant experience [(i.e., base experience)] and the level of variance in the subject experience [(i.e., company experience)]. The GACT method is not always practical in applying credibility because of the type of data that is required to evaluate the 'accuracy' of relevant experience [(i.e., base experience)]. For example, in the case of standard mortality tables, details about the individual contributions (such as company name or plan) to the standard mortality table are generally not publicly available. This information is necessary to evaluate the variability of the mortality rates for the individual contributions relative to the estimated composite rates of death from the standard mortality table (i.e., evaluate the 'accuracy' of relevant [base] experience). Due to the lack of necessary data for a GACT analysis, the LFCT method is usually used in applying credibility to mortality.[42]

The CIA Educational Note makes the following statements about GACT:[43]

> The Greatest Accuracy Credibility Theory (GACT) or 'European credibility' is based on work by Bühlmann. GACT has a better theoretical basis than LFCT, and ensures that results are 'balanced,' so normalization is obviated. Greatest Accuracy Credibility Theory allows one to estimate within and between sub-category sources of variation . . .
>
> GACT is theoretically complete, and meets the criteria for a credibility method with one shortcoming. The shortcoming is that additional information about industry experience (beyond what is customarily collected and published) is required. Without these practical difficulties, GACT would likely be the preferred credibility method to use in determining the expected valuation mortality assumption . . .
>
> From a theoretical point, the GACT method is preferable since it is theoretically complete. However, current industry data is not sufficiently detailed to support the use of GACT.

Given the considerations noted above, and the survey results that indicate GACT is not typically used in the Canadian market, GACT was not considered further in the analysis.

Of interest, under the U.S. principle-based reserving standards related to individual life insurance, which are described in Chapter 20 of the National Association of Insurance Commissioners Valuation Manual, 2019 Edition ("VM-20"), values are provided for use with the GACT approach to approximate the variation between each company's mean and the overall mean, which is a required component of the method.[44] The fact that values are provided for use with the GACT approach facilitates insurers' use of this method.

Limited Fluctuation Credibility Theory

As noted above, both LFCT and GACT approaches use the same linear estimator formula to combine the company experience and the base experience[45] (see Formula 1 above).

The LFCT method is based on confidence intervals.[46] The CIA Educational Note includes the following points about this aspect of the LFCT method:[47]

---

[42] SOA Pension Credibility Educational Resource, Section 2.2.1.
[43] CIA Educational Note, Sections 560 and 570.
[44] SOA *Financial Reporter* PBA Corner Credibility Article, page 23.
[45] CIA Educational Note, Sections 530 and 540, SOA Pension Credibility Educational Resource Section 2.2.
[46] SOA Credibility Theory Practices Report, page I.4.
[47] CIA Educational Note, Section 540.

- In LFCT, one calculates X_E by selecting a range parameter r (r > 0) and a probability level p (0<p<1) such that the difference between X_E and its mean μ is small.[48]
- The criterion can be written as $\Pr\{|X - \mu| \le r\lambda\} \ge p$ [49] where r is the error margin, and p is the confidence level. Parameter values of p = 90% and r = 3% are interpreted as a 90% probability of being correct within a 3% margin of error.
- In other words, X_E is a good estimate of future expected mortality if the difference between X_E and its mean μ is small relative to μ with high probability.

The SOA Pension Credibility Educational Resource makes the point that

> [f]ull credibility is assigned to subject [i.e., company] experience when there is enough subject experience that the error in the estimate is within an acceptable limit with sufficiently high probability. Partial credibility is assigned to subject experience when the variance of the estimate is too high due to lack of data. The definitions of 'acceptable limit' and 'sufficiently high probability' require subjective judgment, so LFCT is not considered as objective as GACT.[50]

The researchers note this is an important consideration for LFCT. Thus, the in-depth study of the LFCT method performed by the researchers involves varying parameters (i.e., using a range of expected errors and confidence levels).

The development of the formulas for LFCT presented in the CIA Educational Note is based on the assumption of a Poisson distribution for the number of observed events (e.g., number of deaths, number of lapses). The CIA Educational Note includes several points about this aspect of the LFCT method that the researchers would like to highlight.

First, the CIA Educational Note makes the following point:[51]

> 1. Although the theoretical distribution for mortality is binomial, when the probabilities of the event (death, represented by the random variable X in the above formulas) are small, the Poisson distribution provides a reasonable approximation to a binomial distribution.

The researchers note that the question of whether or not the probabilities of the event are small enough in order for the Poisson to provide a reasonable approximation to a Binomial distribution is an important consideration and potential limitation of this approach. This aspect of the LFCT method is explored further in the credibility analysis of the lapse sample data sets in Section 4.5.

Second, the CIA Educational Note makes the following point:[52]

> 2. In the Simple Poisson Model, the only random variable is the number of claims, which is assumed to be Poisson.[53] Variations in claim size are ignored. If there is significant dispersion in the net amount at risk for each policy in the block under consideration, the use of a Simple Poisson Model may be inappropriate. The Compound Poisson Model incorporates the effect of variation in claim size, and would normally result in a higher threshold of claims needed to reach the same credibility level. The Compound Poisson Model is discussed in Appendices 1 and 2.

The researchers note that the question of whether or not there is significant dispersion of the "outcomes" (e.g., claim size, net amount at risk) for each policy in the block is an important consideration of the LFCT approach. In the Simple

---

[48] The researchers note that the calculation of X_E assumes the difference between X_E and its mean μ is small *relevant to the mean*.
[49] The CIA Educational Note defines the Poisson parameter to be λ, which would be the mean μ for the simple Poisson model. If X_i is Poisson with parameter λ, a consistent probability criterion is given by the following: $\Pr\{|X_E - \lambda| \le r\lambda\} \ge p$. The researchers also note that a generalized probability criterion is given by the following: $\Pr\{|X_E - \mu| \le r\mu\} \ge p$.
[50] SOA Pension Credibility Education Resource, Section 2.2.2.
[51] CIA Educational Note, page 16, item 1.
[52] CIA Educational Note, page 16, item 2.
[53] The CIA Educational Note includes the following footnote for this statement: "See Loss Models: From Data to Decisions Example 5.20 or Introductory Credibility Theory Example 3.2.2."

Poisson Model, variations in claim size are ignored. The researchers refer to this criterion for defining credibility as "by policy" or "by count." The Compound Poisson Model incorporates the effect of variation in claim size. The researchers refer to this criterion for defining credibility as "by amount." This aspect of the LFCT method is explored further in the credibility analysis of the sample data sets.

Third, the CIA Educational Note makes the following points:[54]

3. Parameter values p = 90% and r = 5% are frequently cited as the minimum levels required for full credibility; however, there is no theoretical basis for determining these parameter values. When setting the expected mortality assumption for valuation purposes, one may want to use a higher threshold for full credibility, such as p = 90% and r = 3%. These parameters were the subject of many discussions within the Task Force and within CLIFR. The consensus was that a minimum of 3,007 deaths would be recommended for 100% credibility. We expect that this issue will be revisited periodically as new literature and research emerges in this area.
4. For p = 90% and r = 3%, the factor for partial credibility is defined by

$$Z = \min\left\{\sqrt{\frac{n}{3007}}, 1\right\} \qquad \text{(Formula 2)}$$

Where n = number of claims in experience data and 3,007 is taken from the standard normal table.[55]

Number of Claims	30	120	271	481	752	1083	1473	1924	2436	3007
Z	0.10	0.20	0.30	0.40	0.50	0.60	0.70	0.80	0.90	1.00

5. The parameters defined in Step 4 above are suggested for use in most situations. A significant dispersion in net amount at risk in the inforce block will increase volatility and could result in the need to use a higher number of deaths.

The number of claims needed for full credibility under other values of p (confidence level) and r (margin of error) are set out in the following Standard Normal Table in Appendix 2 of the CIA Educational Note:[56]

Standard Normal Table – Range and Probability Parameters					
Number of Claims Needed for Full Credibility					
Probability Parameter p	Range Parameter r				
	5%	4%	3%	2%	1%
90%	1,082	1,691	3,007	6,765	27,060
95%	1,537	2,401	4,268	9,604	38,416
99%	2,654	4,147	7,373	16,589	66,538
99.9%	4,331	6,767	12,030	27,068	108,274

[54] CIA Educational Note, page 16, item 3; page 17, items 4 and 5.
[55] The CIA Educational Note includes the following footnote for this statement: "The credibility factors set out in the previous CIA standard VTP 6 Expected Mortality Experience for Individual Insurance were based on LFCT using a simple Poisson distribution. The factors incorporate a conservative bias that depended, in part, upon whether industry or company data is better. Therefore, the credibility factors in VTP 6 are different than those obtained using the above formula. Since the objective is to select the expected valuation assumption, a conservative bias is not appropriate."
[56] CIA Educational Note, Appendix 2, page 38.

The researchers make the following observations with respect to these three points:

- For the LFCT (Simple Poisson Model, "by count" or "by policy"), the formula for the number of claims needed for full credibility is:[57]

  $$\text{Number of Claims Needed for Full Credibility} = \frac{z^2}{r^2}$$

  where

  - z is the critical value for a standard normal random variable for a given confidence level (the probability parameter, p); and
  - r is the margin of error which is referred to in the CIA Educational Note as the "Range Parameter."

- The CIA Educational Note recommends 3,007 deaths for full credibility (with the caveat that dispersion in net amount at risk and the absence of credible industry data are two significant factors that would be considered when determining the number of deaths needed for full credibility).[58] Using the formula noted above, this result may be computed as follows:

  $$\text{Number of Claims Needed for Full Credibility} = \frac{z^2}{r^2} = \frac{1.645^2}{.03^2} = 3{,}007$$

  where

  - 1.645 is the critical value for a standard normal random variable for a 90% confidence level; and
  - r is the margin of error or "Range Parameter" of 3%.

  (Other results in the table are derived in a similar manner.)

- In the survey responses, several companies reported using 3,007 deaths for valuation purposes for full credibility. However, one company reported using 6,014 deaths. The company using 6,014 deaths explained that using 3,007 deaths assumes policy counts (not amounts, which may exhibit more variability) and may understate true full credibility. The company uses 6,014 deaths in order to recognize variability in coverages and age distribution. Other companies echoed the concern that the 3,007 number would not be adequate when comparing non-homogenous lives.
- The SOA Pension Credibility Educational Resource indicates that the proposed Internal Revenue Service regulations define full credibility as 1,082 deaths, which is based on a 5% margin of error with a 90% confidence level. It also demonstrates how the generalized formula (shown above) and the result itself are derived.[59]
- Section 9.C.4 of VM-20 requires the margin of error to be at most 5% and the confidence level of at least 95%. In addition, the A/E mortality ratios are required to be calculated "by amount."
- The credibility analysis for mortality focuses on a comparison of the number of deaths needed for full credibility "by count" and "by amount" for a range of confidence levels/margins of error, products and risk classifications.
- The credibility analysis for lapse focuses on a comparison of the number of lapses needed for full credibility "by count" and "by amount" for a range of confidence levels/margins of error and certain products.

---

[57] The tables produced by the researchers indicate that the number of claims needed for full credibility for p=99% and r=1% is 66,358. The researchers note that for LFCT "by count" the number of claims needed for full credibility is equal to z^2/r^2. For a probability parameter of 99%, z=2.576. Thus, $z^2/r^2 = 2.576^2/.01^2 = 66{,}358$. The researchers also note that the z values used in deriving the results in the CIA Educational Note are rounded.

[58] CIA Educational Note, Section 570, item 3.

[59] SOA Pension Credibility Educational Resource, Section 2.3.

The LFCT method only uses data from the particular company being studied to determine the credibility factor.[60] Thus, the LFCT assumes that the base experience accurately represents the quantity it is estimating.[61] The researchers note that this is an important consideration and potential limitation of this approach. Some survey respondents noted the lack of appropriate base experience as an issue. The CIA Educational Note also highlights the importance of this potential limitation:[62]

> The use of the blending methodology set out in this section assumes that there exists relevant industry basis for blending. If there is no industry table or study that corresponds to the company's business mix, then it may be appropriate to assign a higher credibility factor to the company data than otherwise.

The analysis does not explore this potential limitation of the method (i.e., that appropriate base experience is not available). However, the survey results include commentary about this point.

Finally, the CIA Educational Note makes the following points:[63]

- The Poisson application can be extended to include data from more than one period or year. However, the number of years would be limited so that the mix and material risk characteristics of the portfolio are homogeneous over time.
- Application of LFCT to Sub-categories of Business
    1. If the actuary wants to reflect experience split by sub-category (perhaps by sex, product, or duration) but the experience in those sub-categories is not 100% credible, the actuary would decide either to use the overall credibility factor, or the lower credibility for the amount of experience in that sub-category.
    2. One can pool disparate distributions within the aggregate data under certain conditions. Several approaches are discussed in Appendix 2.

The credibility analysis does not address including data for more than one period or year or the application of LFCT to sub-categories of business. However, various approaches to balancing the need for sufficient data (longer time periods) and recent data (shorter time periods) were described by survey respondents. Also, some of the respondents noted that they use the Normalized Method to calculate credibility factors by sub-category.

The Normalized Method (which is a variant of LFCT used to reflect experience split by sub-category) is summarized in the CIA Educational Note and is described as the favored approach because, except for its theoretical shortcomings, it meets all of the criteria for a good credibility method.[64] The CIA Educational Note includes the following as desirable characteristics for a good credibility method: the method is practical to apply; the sum of expected claims for the within-company sub-categories is equal to the total company expected claims; all of the relevant information is used; the results are reasonable in extreme or limiting cases; and sub-category A/E ratios are reasonable relative to company and industry data (e.g., they fall within the range of corresponding industry and company experience A/E ratios).[65]

---

[60] SOA Credibility Theory Practices Report, page I.4.
[61] SOA Pension Credibility Educational Resource, Section 2.2.2.
[62] CIA Educational Note, page 17, item 6.
[63] CIA Educational Note, page 17.
[64] CIA Educational Note, page 21, Section 570, item 2.
[65] CIA Educational Note, pages 14–15, Section 520.

About the Canadian Institute of Actuaries

The CIA is the national, bilingual organization and voice of the actuarial profession in Canada. Our members are dedicated to providing actuarial services and advice of the highest quality. The Institute holds the duty of the profession to the public above the needs of the profession and its members.

The Head Office has a dedicated group of 30 staff members located in Ottawa. The Head Office looks after publications, communications, member services, translation, volunteer support, maintaining the website and professional development.

The CIA Board has 15 actuaries, six councils focused on the core needs of the profession, and over 40 committees and numerous task forces working on issues linked to the CIA's strategic plan.

The CIA

- Promotes the advancement of actuarial science through research;
- Provides for the education and qualification of members and prospective members;
- Ensures that actuarial services its members provide meet extremely high professional standards;
- Is self-regulating and enforces rules of professional conduct; and
- Is an advocate for the profession with governments and the public in the development of public policy.

CIA Head Office
Canadian Institute of Actuaries
360 Albert Street, Suite 1740
Ottawa, Ontario K1R 7X7
www.cia-ica.ca

About the Society of Actuaries

The SOA, formed in 1949, is one of the largest actuarial professional organizations in the world and is dedicated to serving more than 32,000 actuarial members and the public in the United States, Canada and worldwide. In line with the SOA Vision Statement, actuaries act as business leaders who develop and use mathematical models to measure and manage risk in support of financial security for individuals, organizations and the public.

The SOA supports actuaries and advances knowledge through research and education. As part of its work, the SOA seeks to inform public policy development and public understanding through research. The SOA aspires to be a trusted source of objective, data-driven research and analysis with an actuarial perspective for its members, industry, policymakers and the public. This distinct perspective comes from the SOA as an association of actuaries, who have a rigorous formal education and direct experience as practitioners as they perform applied research. The SOA also welcomes the opportunity to partner with other organizations in our work where appropriate.

The SOA has a history of working with public policymakers and regulators in developing historical experience studies and projection techniques as well as individual reports on health care, retirement and other topics. The SOA's research is intended to aid the work of policymakers and regulators and follow certain core principles:

Objectivity: The SOA's research informs and provides analysis that can be relied upon by other individuals or organizations involved in public policy discussions. The SOA does not take advocacy positions or lobby specific policy proposals.

Quality: The SOA aspires to the highest ethical and quality standards in all of its research and analysis. Our research process is overseen by experienced actuaries and non-actuaries from a range of industry sectors and organizations. A rigorous peer-review process ensures the quality and integrity of our work.

Relevance: The SOA provides timely research on public policy issues. Our research advances actuarial knowledge while providing critical insights on key policy issues, and thereby provides value to stakeholders and decision makers.

Quantification: The SOA leverages the diverse skill sets of actuaries to provide research and findings that are driven by the best available data and methods. Actuaries use detailed modeling to analyze financial risk and provide distinct insight and quantification. Further, actuarial standards require transparency and the disclosure of the assumptions and analytic approach underlying the work.

Society of Actuaries
475 N. Martingale Road, Suite 600
Schaumburg, Illinois 60173
www.SOA.org