Credibility Educational Resource for Pension Actuaries

Application of Credibility Theory to Mortality Assumption

August 2017

Credibility Educational Resource for Pension Actuaries

Application of Credibility Theory to Mortality Assumption

AUTHOR Irina Pogrebivsky, FSA, EA

SPONSOR Society of Actuaries (SOA)
Pension Section Council

The development of this paper was sponsored by the SOA Pension Section Council. The Council would like to acknowledge the work of the primary author, Irina Pogrebivsky, FSA, EA, along with the significant input of the project oversight group:

- Julie Curtis, FSA, EA, MAAA (Chair)
- Gavin Benjamin, FSA, FCIA
- David Kausch, FSA, EA, FCA, MAAA, MSPA
- Judy Ocaya, FSA, EA, FCA
- Faisal Siddiqi, FSA, FCIA
- Aaron Weindling, FSA, EA, FCA, MAAA
- Cynthia MacDonald, FSA, MAAA, SOA senior experience studies actuary
- Andrew Peterson, FSA, EA, FCA, MAAA, SOA senior staff fellow, Retirement Systems

Additional support and review assistance was provided by Stuart Klugman, FSA, CERA, SOA senior staff fellow, Education; and Mervyn Kopinsky, FSA, EA, SOA experience studies actuary.

CONTENTS

1 Executive Summary

1.1 Background

In recent years, the pension community has focused increased attention on the mortality assumptions used for valuing pension plans. This has been driven by the confluence of several factors:

- The recent publication of new pension mortality tables in both the United States and Canada—for example, the RP-2014/MP-2014/MP-2015/MP-2016 tables/scales by the Society of Actuaries (SOA) in the United States and the CPM 2014 tables and new improvement scales by the Canadian Institute of Actuaries (CIA) in Canada
- The increased focus on pension de-risking and pension risk transfer transactions and the impact of mortality table differences on pricing transactions
- A general awareness of the need to better understand and reflect emerging longevity trends

With more attention on mortality assumptions, there is also a growing awareness of the variability of mortality within different demographic groups and/or plan populations resulting in a desire for more customized assumptions. For example, evidence suggests that mortality varies by industry, geography and job type (i.e., collar). In addition, the size of pension benefit amounts can be a predictor of mortality, and certain plan provisions may lead to anti-selection, which affects future mortality (such as the ability of former plan members to elect a lump sum). This has led to a renewed interest in actuarial credibility theory as a means for adjusting standard tables to better fit specific pension plan populations in both the United States and Canada.

Additionally, in the U.S. context, the Bipartisan Budget Act of 2015 included a provision that required the IRS to issue guidance that would allow plans to reflect their own mortality experience rather than use the IRS-prescribed mortality tables in “accordance with established actuarial credibility theory.” Accordingly, in December 2016, the IRS issued proposed regulations (*proposed IRS mortality regulations*) outlining the requirements a plan must meet in order to obtain IRS approval to use plan-specific mortality tables (which includes guidance on the use of credibility procedures) for purposes of pension plan valuations. This development, along with the expansion of Actuarial Standard of Practice No. 25 (Credibility Procedures) to cover pension work, has further increased the interest in credibility theory for U.S. pension actuaries.

1.2 Objectives

While the use of credibility theory has been standard practice among insurance actuaries, its use by pension actuaries has traditionally been limited.

Credibility is a mechanism for estimating a quantity of interest by combining subject experience and relevant experience to improve the estimate. *Subject experience* is the actual experience we are trying to estimate, whereas *relevant experience* is derived from other related sources and can be used to estimate the quantity of interest. Subject experience is useful because it is directly related to the quantity of interest. Relevant experience is usually necessary because subject experience often lacks sufficient volume to provide the desired level of "accuracy."

The objectives of this paper are as follows:

- To provide a clear and concise overview of credibility theory
- To describe and compare the standard approaches
- To provide guidance to pension actuaries on how to apply the theory with respect to mortality
- To showcase examples of real-life situations that can be used as a basis for evaluating relevant application based on a specific scenario
- To provide a list of resources on credibility theory that exist within the actuarial profession

1.3 Limitations of Use

- This paper will focus on *base mortality table* selection, adjustment and/or development. Base mortality reflects the mortality rates of plan members at the time of the mortality experience study.
- Future mortality improvements are outside the scope of credibility analysis and application described in this paper. It is assumed that projection scales, reflecting future mortality improvements, are applied to the base table *after* adjustments for subject experience have been incorporated.
- This paper does not reflect any rules on the selection of mortality tables that may be imposed by applicable legislation or regulators. Also it is not the intent of this paper to provide an opinion on the appropriateness of any specific method mentioned for selecting, adjusting and/or developing a mortality table. This paper is not and should not be treated as a standard of practice or a practice note; it is intended solely as an educational resource for practitioners. The SOA makes no warranty, guarantee or representation, either expressed or implied, regarding this paper. In no event shall the SOA be liable for any damages in connection with or resulting from reliance on any aspect of this paper. The user assumes all responsibility for the use of information contained in this paper.

2 Overview of Credibility Theory

2.1 What is Credibility Theory?

Actuaries have long been challenged with incorporating experience data into the evaluation of a company's overall risk and financial condition as well as estimating reserves and cost of benefit plans. Actuarial judgment based on experience studies is the predominant method currently employed. Credibility procedures can validate or improve the actuarial judgment applied to an organization's experience data.

Credibility procedures use statistical approaches to adjust relevant experience-based assumptions. As already mentioned, credibility attempts to combine subject experience and relevant experience to improve the estimate of the desired quantity of interest.

Life and property/casualty actuaries have traditionally used credibility in setting premiums by modifying class rates (relevant experience) to reflect experience (subject experience). *Classes* are groupings of risks with similar characteristics, although each risk still has its own unique characteristics. An insurance premium charged to each risk starts with a rate common to the class. This rate is then usually adjusted to reflect the individual's experience so the premium for that individual risk is based on both the class rate and past loss experience for the individual. The adjustment depends on how much "credibility" is assigned to the experience of the individual risk.

Example: Auto rate pricing for Jack and Jill from Hill Insurance Company

- Jack has had 2 accidents over the last 5 years (0.4 accidents per year on average).
- Jill has had 0 accidents over the last 5 years (0 accidents per year on average).
- Company experience for *all* insured Hill drivers is an average of 0.05 accidents per year.

If we were to fully incorporate Jack and Jill's actual experience, Jack's rate would be 8 times the average rate, and Jill's rate would be 0. This is obviously unfair since Jack may have just had some bad luck but is actually a relatively good driver overall. On the other hand, Jill, although likely a good driver, should pay something for her coverage, as it is unreasonable to assume there is no chance that she will ever have an accident.

Assigning the average rate for the class of *all* Hill drivers of 0.05 to both Jack and Jill is also not optimal, since it is clearly not the best estimate of future experience—Jill has demonstrated that she is likely to be a better driver than Jack. To reflect the best future estimate, Hill should incorporate some of the experience *for each specific driver* into the rates. The question is how much credit (or credibility) should be given to their actual experience.

In evaluating how much credibility to assign to subject experience, consideration should be given to two factors:

1. *The "accuracy" of the* ***relevant experience*** *(such as average overall class rate).* The "accuracy" of relevant experience is measured in terms of how much the expected individual risk's outcomes vary from the expected outcome for the class (the average overall class rate). If there is a small amount of variability (i.e., the overall class rate is very likely to be representative of individual outcomes within the class), then there is a high degree of "accuracy" of relevant experience, and we can assign more weight to it and less weight to subject experience. If, however, relevant experience is less likely to be representative of the individual outcomes (i.e., there is a lot of variability), then we should assign less weight to relevant experience and more weight to subject experience.

2. *The variability of the* ***subject experience*** *(such as individual car driver experience).* A large amount of expected variation in subject experience may indicate that that experience will not be very useful in estimating the expected value; therefore, we should assign less weight (or credibility) to it. A small amount of expected variation in subject experience indicates that we should assign a large amount of credibility to it.

These factors can be broadened to encompass adjustments for actual experience in a pension valuation mortality table context as follows. Credibility is a way of combining relevant experience (estimated rate of death from a standard mortality table) and subject experience (estimated rate of death based on experience from the specific plan's population). As already described, the amount of weight given to each depends on (1) the "accuracy" of the relevant experience and (2) the volatility of subject experience (the plan's observed rate of death). Large variability in subject experience implies less reliable estimates and therefore less credibility. Generally, more subject experience data result in more accurate estimates, less variability and thus higher credibility.

2.2 Types of Credibility Approaches

There are two main approaches to credibility: *Greatest Accuracy Credibility Theory* (*GACT*) and *Limited Fluctuation Credibility Theory* (*LFCT*). Both approaches strive to produce improved estimates of future events based on combining relevant and subject experience.

Both methods use the following linear estimator formula to combine the relevant and subject experience:

$$\hat{E} = Z \cdot \widehat{m} + (1 - Z) \cdot \hat{a},$$

where $\widehat{m}$ is the estimated value of the unknown quantity of interest based on the subject experience, Z is the weight (or credibility factor) assigned to that estimate and $\hat{a}$ is the estimate of that same quantity based on the relevant experience. (Note that $\hat{a}$ is the estimate that would be used if there were no subject experience.)

The difference between the two methods is how *Z* is determined. 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.

Due to the practical limitations of using GACT in the context of mortality adjustments, this paper does not provide a detailed description of that theoretical model. A brief overview of GACT follows, and a more detailed summary of the method is included in the Appendix. This paper focuses on LFCT and its application to the mortality assumption in pension plans. We will first provide a brief overview of LFCT, with a comparison to GACT, and then a more detailed explanation of the method and its application to mortality.

2.2.1 Greatest Accuracy Credibility Theory Overview

GACT is also known as *Bühlmann Credibility* and *Linear Bayesian Credibility*. 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 and the level of variance in the subject 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. 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 experience). Due to the lack of necessary data for a GACT analysis, the LFCT method is usually used in applying credibility to mortality.

2.2.2 Limited Fluctuation Credibility Theory Overview

LFCT is based on a similar linear formula model to that for GACT, but unlike GACT, it does not consider the "accuracy" of relevant experience. The LFCT model looks only at the variability of the subject experience and assumes that relevant experience accurately represents the quantity it is estimating. All attention is on the amount of variation in the subject experience. *Full credibility* is assigned to subject experience when there is enough subject experience data 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. In other words, when the variance of subject experience is not within an acceptable limit, the credibility given to that experience is reduced, and some weight is assigned to relevant experience (which in this model is assumed to accurately represent the quantity of interest). In this way, the expected error in the estimate is reduced to an acceptable range. The definitions of "acceptable limit" and "sufficiently high probability" require subjective judgment, so LFCT is not considered as objective as GACT.

2.2.3 Comparison of GACT and LFCT

Feature	GACT	LFCT
Estimator	$\hat{E} = Z \cdot \widehat{m} + (1 - Z) \cdot \hat{a}$	$\hat{E} = Z \cdot \widehat{m} + (1 - Z) \cdot \hat{a}$
Method	Determine *Z* such that the expected value of the square of the difference between the estimate and the quantity being estimated is minimized.	Full credibility is assigned to subject experience (i.e., *Z* = 1) when there is enough data that the error of subject experience is within an acceptable limit with sufficiently high probability.
Relevant experience	Data required.	No data required. Relevant experience is assumed to be accurate with regard to the quantity it is estimating.
Subject experience	Data required.	Data required.
Degree of statistical rigor	More rigorous—the variability of both relevant experience and subject experience are determined and factored into weighting.	Less rigorous—this model assumes relevant experience accurately reflects the quantity being estimated and requires subjective judgment in determining the credibility weighting.

2.3 Limited Fluctuation Credibility Theory

As stated earlier, the LFCT estimator is $\hat{E} = Z \cdot \widehat{m} + (1 - Z) \cdot \hat{a}$, where $\widehat{m}$ is the estimated value of the unknown quantity of interest based on the subject experience, Z is the weight (or credibility factor) assigned to that estimate and $\hat{a}$ is the estimate of that same quantity based on the relevant experience. (Note that $\hat{a}$ is the estimate that would be used if there were no subject experience.) In the context of insurance premiums, $\hat{a}$ is the average accident rate for the class of drivers (such as 0.05 for Hill Insurance from the earlier example).

In this model, *Z* is set to 1 if there is a high probability of being within a very small margin of relative error with respect to the true value: $\Pr(|\widehat{m} - m| \leq rm/100) \geq p$, where $\hat{m}$ is the estimate based on the subject experience, and *m* is the true, unknown value. The choice of confidence level (*p*) and margin of error (*r*) is subject to judgment. This is one of the main disadvantages of LFCT—it doesn't provide an objective basis for determining full credibility. Subjective judgment is required to determine the quantity of data necessary for this determination. A 90% confidence level and 5% margin of error (*p* = 90% and *r* = 5%) are frequently cited as minimum levels required for full credibility; however, since there is no theoretical basis for this threshold, other assumptions may be just as valid. For example, a CIA educational note titled "Expected Mortality: Fully Underwritten Canadian Individual Life Insurance Policies" (described in Section 2.4.3) recommends using a 3% margin of error with a 90% confidence level for purposes of setting mortality assumption in valuations for insurance practitioners; implying credibility is achieved at 3,007 deaths. On the other hand, the proposed IRS mortality regulations define full credibility at 1,082 deaths, which is based on a 5% margin of error with a 90% confidence level, as described here.

In the context of mortality at a specific age (x), it turns out that 1,082 observed deaths are required for *fully credible* subject experience data using a 90% confidence level and a 5% margin of error (p = 90% and r = 5%). The derivation is as follows:

For notation, let N be the number of observed lives at a given age (x), let d be the observed number of deaths and let q be the true mortality probability. The estimate of the quantity of interest (q) is then d/N, and full credibility is assigned if the following is true:

$$\Pr\left(\left|\frac{d}{N}-q\right| \le 0.05q\right) \ge .9$$

To solve this, we need an underlying distribution for the actual deaths. If we assume that all lives are exposed for the full year, all lives have the same value of q and their deaths are independent, the actual number of deaths will have a binomial distribution. However, if N is large enough, the binomial distribution can be approximated by a normal distribution, with mean Nq and variance $Nq(1-q)$. Then we have

$$\Pr\left(\left|\frac{d}{N}-q\right| \le 0.05q\right) \ge 0.9$$

$$\Pr(0.95Nq \le d \le 1.05Nq) \ge 0.9$$

$$\Pr\left(\frac{0.95Nq-Nq}{\sqrt{Nq(1-q)}} \le z \le \frac{1.05Nq-Nq}{\sqrt{Nq(1-q)}}\right) \ge 0.9$$

Here z represents a standard normal random variable. This equation can be simplified by noting that $1-q$ is essentially equal to 1. Making this change yields

$$\Pr\left(-0.05\sqrt{Nq} \le z \le 0.05\sqrt{Nq}\right) \ge 0.9$$

For a standard normal random variable, $\Pr(-1.645 \le z \le 1.645) = 0.9,$ so the equation is satisfied when $0.05\sqrt{Nq} \ge 1.645$ or $Nq \ge 1{,}082.$ Thus, the expected number of deaths must be at least 1,082 for the full credibility requirement to be met. Because the number of expected deaths is not known, the actual number of deaths is used when determining if there is full credibility.

Generalizing this for any p and r leads to the conclusion that full credibility is achieved when the observed number of deaths is greater than or equal to $(z_{(1+p)/2} / r)^2$, where $z_{(1+p)/2}$ is the given percentile from the standard normal distribution.

Note that this model was developed based on number of deaths rather than benefit amounts. Thus, this solution is based on a counts-weighted approach to developing mortality assumptions. Models for an amounts-weighted approach are discussed later in this paper.

To the extent that there is not enough subject experience data for full credibility (i.e., $Z < 1$), some portion of credibility needs to be assigned to the relevant data as well. Lack of full credibility implies that the variance of the estimator is too high and needs to be reduced. Thus, *Z* is set in such a way that the variance of the estimator is equivalent to the variance when there is enough data for full credibility.

When there is exactly enough data for full credibility, $Z = 1$ and

$$Var(\hat{E}) = Var(\hat{m}) = Var(d/N) = \frac{q(1-q)}{N_f} \approx \frac{q}{N_f},$$

where N_f is the exposure for full credibility. Note that the equation is simplified since $1 - q$ is essentially equal to 1.

When *N* is too small, the variance will be too large. But applying the credibility formula (and recalling that the estimate based on the relevant experience is not random for LFCT) reduces the variance:

$$Var(\hat{E}) = Var\,(Z \cdot \hat{m} + (1 - Z) \cdot \hat{a}) = Z^2 Var\left(\frac{d}{N}\right) \approx \frac{Z^2 q}{N_a},$$

where N_a is the actual exposure.

Matching the two variances

$$\left(\frac{Z^2 q}{N_a} = \frac{q}{N_f}\right)$$

leads to $Z = \sqrt{N_a q / N_f q}$. This means the credibility factor is the square root of the ratio of the observed deaths (the expected deaths are not available) to the number of deaths required for full credibility.

LFCT is less rigorous than GACT. The model is only designed to ensure that the error around the subject experience data is minimized to an acceptable level. Subjective judgment is required regarding the acceptable level of error. The setting of the *r*, *p*, and *a* parameters is subjective, and LFCT does not account for variances (or errors) in the relevant data. However, in situations where relevant experience can be assumed to be “correct” (as in the case of standard mortality tables), this is a very useful method because it does not require assessment of the variance of the relevant experience data and is a practical approach.

2.4 Application of Actuarial Standards of Practice

Warning: The guidance that follows is current as of the publication of this paper in August 2017. Any additional or revised guidance should be taken into account as of the date of the analysis.

2.4.1 Actuarial Standard of Practice No. 25

(December 2013)

Actuarial Standard of Practice (ASOP) No. 25, Credibility Procedures, provides guidance to U.S. actuaries in selecting and developing credibility procedures. Procedures covered by the standard include "(1) evaluating subject experience for potential use in setting assumptions without reference to other data and (2) to improve the estimate of the parameter under study."

ASOP No. 25 does not provide a specific recommendation for how to apply credibility procedures, but it does provide general direction on the selection and development of the procedure to be used, the selection of the relevant experience to be blended and the professional judgment to be used in assigning credibility to the subject experience. The standard has precise and helpful definitions of credibility terminology, much of which has been used in this paper. It also provides guidance on appropriate disclosures when communicating results based on information developed by using credibility procedures.

The Appendix of ASOP No. 25 refers to LFCT and GACT, among other approaches, but does not recommend a particular approach.

2.4.2 Actuarial Standard of Practice No. 35

(September 2014)

ASOP No. 35, Selection of Demographic and Other Noneconomic Assumptions for Measuring Pension Obligations, provides guidance to U.S. actuaries on selecting the demographic assumptions and adjustments necessary to customize for a plan's particular population demographics. For example, the standard suggests the following considerations for selecting demographic assumptions:

- The characteristics of the covered group
- Any features of or change in the plan design that may influence the assumption
- Appropriate experience from the specific plan and other relevant sources
- Specific experience of the covered group or other groups with similar characteristics that may be useful in forming a judgment about future expectations

Specifically for mortality, ASOP No. 35 also recommends that actuaries should consider "the use of different assumptions for different participant subgroups and beneficiaries."

To the extent that the listed considerations warrant modifications to the standard mortality tables based on subject experience, credibility methods may be helpful in performing the adjustments.

2.4.3 Canadian Institute of Actuaries Standards of Practice and Educational Notes

Section 1730 of the Standards of Practice provides guidance to Canadian actuaries on the use of appropriate assumptions. In addition, two educational notes provide comprehensive direction on applying credibility methods and selecting mortality assumptions:

- “Expected Mortality: Fully Underwritten Canadian Individual Life Insurance Policies” (July 2002, Document 202037) describes methods for blending company and standard experience using credibility methods. This note describes both LFCT and GACT.
- “Selection of Mortality Assumptions for Pension Plan Actuarial Valuations” (March 2014, Document 214029) helps Canadian actuaries choose mortality assumptions. The educational note specifically refers to the use of credibility in reflecting actual plan experience: “The first step in developing an appropriate best estimate mortality assumption is to determine the best estimate of the current levels of mortality. The best estimate would be developed considering the plan’s actual plan mortality experience (where available), the credibility of such plan experience, the experience of similar plans, published mortality studies and possible adjustments based on characteristics such as collar type, industry and pension size.”

3 Application of Credibility Theory to Mortality Assumption in Pension Plan Valuations

Standard mortality tables reflect a large cross-section of populations across many pension plans. Of course, each specific individual and/or smaller grouping of individuals may have a different expected rate of death within that broad population. Credibility procedures are used to modify a broad population table to reflect the experience of a subgroup for which a more appropriate estimate of mortality can be determined by incorporating that subgroup's own experience. For pension plan valuations, the goal is to incorporate the experience of the participants covered by the plan being valued.

3.1 Data Needed for Credibility Analysis

Specific company mortality data provide the experience-based results needed for weighting with the relevant experience. Data collected should reflect the specific population or group for whom the assumption is being set. Generally, the mortality assumption that is most critical in pension valuations is the one that affects participants currently or projected to be in pay status. So it is common to isolate data collection to those already receiving a benefit from a specific employer plan. Note that an employer may have more than one plan, and generally the mortality assumption is plan-specific. The actuary should consider comparing mortality experience between plans to validate whether separate adjustments should be made. Furthermore, a single plan may cover several subgroups of employees (such as white-collar and blue-collar workers). If the actuary believes that mortality rates are consistent for all plan participants, a single mortality experience study can be performed. However, if there is reason to believe that different mortality may be applicable to each subgroup within a plan, separate studies are needed. In general, actuarial judgement should be used to decide whether different mortality adjustments (and, thus, separate mortality studies) are required for different subgroups. The proposed IRS mortality regulations stipulate that in constructing a plan-specific mortality table, mortality adjustments should be based on the specific experience for each subgroup covered by the plan.

It is highly recommended that all actuaries consider splitting plan experience by gender. There is ample evidence to suggest that mortality experience between males and females is significantly different. U.S. funding regulations require that gender-specific standard mortality tables be used for pension valuations. In addition, the current and proposed regulations stipulate that if plan-specific mortality is used, it must be based on gender-specific experience data. Typically, gender information is readily available, so information should be collected by gender.

To do a mortality experience study, identify the group of retirees and beneficiaries in one plan or across multiple plans. Then for each group, collect the following for each year of the study period:

1. ID code (so each retiree can be tracked individually)
2. Date of birth (or age at measurement date)
3. Date of death (if applicable)
4. Gender
5. Benefit amount (see Section 3.4.3 for the use of amounts- versus counts-weighting in credibility modeling)

The actuary should consider the number of years of experience to include in the study. The assumption must be set as of the measurement date; therefore, data that are too far away from the measurement date should not be used. However, it is beneficial to have data that span multiple years because (1) it increases the amount of exposures for greater credibility, and (2) it lowers the risk of using an anomalous year of data, which can skew results in an inappropriate manner. Generally, three to five years is a good rule of thumb to use in an experience study.

The actuary should also consider the ages of retirees or beneficiaries included in the study. For traditional plans with clearly defined retirement eligibility, the group included in the study will, for the most part, automatically be limited by the age of earliest retirement eligibility (usually 50 to 55 and above). However, for cash balance plans where individuals can "retire" after three or more years of service regardless of age, actuaries will need to make judgment calls on the age range of the data to incorporate into their experience studies. The proposed IRS mortality regulations only allow experience studies that include ages 50 and above to apply to the entire population (even when the study includes individuals younger than 50).

3.2 Building a Mortality Table From Scratch

To build a mortality table from scratch, an actuary would need to estimate q_x at each age (or at small age bands) based on the plan's experience. A tremendous amount of data would be required for the number of deaths in the study to be large enough to achieve a reasonable amount of, if not full, credibility. Section 2.3 showed that under LFCT, 1,082 deaths are required at each age to achieve full credibility on a counts basis (assuming r = 5% and p = 90%). Since the probability of death is small at most ages, the amount of experience needed is typically very large. For example, if q_{75} = .025, 1,082/0.025 = 43,280 lives at age 75 are required to achieve full of credibility for the mortality rate at that age. This is a huge amount of data, especially since it will generally need to be separated by gender.

In addition, to build a new table from scratch, rates will need to be adjusted to create a smooth table. Although this can be accomplished using graduation techniques, extra effort and judgment are required.

Thus, for most organizations, building a table from scratch is impractical. A more palatable approach is to take an existing standard mortality table and adjust it using LFCT methodology. The remainder of this section focuses on this approach.

3.3 Selection of Standard Mortality Table for Credibility Analysis

It is important for the actuary to be thoughtful when selecting the appropriate relevant experience (i.e., standard valuation mortality table) to blend with the subject experience, since the standard table selected will affect the shape of the blended table, and the subject experience may not be fully credible. Many standard mortality tables have been developed that can be used as a basis. Generally, it is preferable to use more recently published tables, unless there is a specific reason to use historical tables.

Aside from the relative age of the table, the actuary should do the following when selecting an appropriate standard table:

- Consider the shape of the table compared to the shape of the actual experience (see Section 3.5.2 for more information on the importance of shape in the mortality credibility analyses).
- Factor in as many of the group's specific characteristics as are available in the standard tables. For example, if the actuary is developing a female table, the base table should be female. If the plan covers blue-collar workers, then consideration should be given to reflecting blue-collar adjustments, if available

for the standard tables (the same is true for white-collar plans). Other characteristics reflected in some existing standard tables include healthy versus disabled annuitants as well as pre-retirement versus post-retirement tables and adjustments for industry.

The base table should be adjusted with appropriate standard mortality improvements during the study period. One possible approach for such an adjustment would be to project the base table with mortality improvements to the midpoint of the experience study so that expected deaths from the standard table are consistent with the actual deaths in the experience study. Other approaches for such an adjustment are also possible.

See the Appendix for a list of existing SOA and CIA standard tables as of the publication date of this paper.

3.3.1 Generational Mortality

A generational mortality table consists of a base table and a multidimensional mortality improvement projection. Thus, credibility procedures should be applied to the base table prior to applying generational projections. In other words, as with nongenerational mortality, the base table should be adjusted for mortality improvement during the study period. One approach for such an adjustment would be to project the base table with mortality improvements to the midpoint of the study period; however, other approaches for such an adjustment are also possible. This adjusted base table should then be modified based on plan experience using credibility procedures as discussed in the rest of this section. Future generational improvement projections should then be applied to that modified base table.

3.4 Adjusting a Standard Table to Reflect Plan Experience

As discussed earlier, GACT is not practical to use in adjusting mortality assumptions, because not enough information about the "accuracy" of the relevant experience (i.e., standard mortality tables) is currently available. Therefore, LFCT is the method generally used to adjust the mortality assumption for plan experience.

The 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 that 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.

So far, the discussion of the LFCT model as it applies to mortality has focused on the estimator of q_x at a specific individual age (x). Now the focus of the model turns to the estimator of the *multiple that will shift the entire mortality table*.

Based on the assumption that a multiple of the standard mortality table will result in the "true" mortality table, the estimator ($\hat{m}$) becomes the factor ($\hat{f}$) that is the standard adjustment to the standard mortality rate at every age (q_x^s). This $\hat{f}$ is defined as the ratio of actual deaths to expected deaths for all ages in the subject experience data. Since $\hat{f}$ is applied at every age (meaning to the whole table) rather than at individual ages, the ratio is determined on an aggregated basis—it reflects summation across all ages.

Full credibility is then achieved when $\hat{f}$ is within the margin of error (*r*) of the "true" *f*, with a probability of at least *p*:

$$\Pr(|\hat{f} - f| \leq rf) \geq p$$

Following this, the LFCT model for adjusting the standard table is constructed as follows (where expected deaths are calculated using the standard mortality table):

$$\hat{f} = \frac{\sum \text{actual deaths}}{\sum \text{expected deaths}}$$

For the formulas that follow, let q_x^s be the standard table mortality rate at age *x* and q_x^F be the estimated mortality rate from the plan experience. The subject experience is $\hat{f} \cdot q_x^s$, and the relevant experience is q_x^s. Then,

$$q_x^F = Z \cdot [\hat{f} \cdot q_x^s] + [1 - Z] \cdot q_x^s = [Z \cdot (\hat{f}) + (1 - Z) \cdot 1] \cdot q_x^s$$

So at any age, the final adjusted "true" mortality rate is a constant multiple of standard rate, where

$$\text{Multiple} = Z \cdot (\hat{f}) + (1 - Z) \cdot 1$$

$$q_x^F = \text{Multiple} \cdot q_x^s$$

The Appendix shows a detailed development of the LFCT model with respect to $\hat{f}$. The next two subsections summarize the results and provide examples of how the adjustment is calculated. A link to the spreadsheet showing how the information in the examples was developed is provided with each example.

3.4.1 Full Credibility

Notation:

i is the *i*th life out of *n* lives studied.

q_i^s is the standard table mortality rate for the *i*th life.

Note that the subscript is not the age. The value of q *is found by looking at that life's age and then applying the value from the standard table.*

d_i is 0 if the *i*th life survives the year of observation and 1 if the life dies.

b_i is the amount associated with the *i*th life.

Then the following quantities can be calculated:

$A_D = \sum_{i=1}^{n} b_i d_i$ = actual death amounts

$A_N = \sum_{i=1}^{n} d_i$ = actual death lives

$E_D = \sum_{i=1}^{n} b_i q_i^s$ = expected death amounts

$E_N = \sum_{i=1}^{n} q_i^s$ = expected death lives

$\lambda = (z_{(1+p)/2}/r)^2$

$\hat{f} = A/E$ = observed mortality ratio

If experience is fully credible, $Z = 1$, and Multiple is simply equal to $\hat{f} = \frac{A_N}{E_N} = \frac{\sum \text{actual deaths}}{\sum \text{expected deaths}}$ from the experience study.

From Section 5.2.1:

Formula 1 (counts): Full credibility ($Z = 1$) on a counts-weighted basis is achieved when

Actual number of deaths $= \lambda = (z_{(1+p)/2}/r)^2$

For a 90% confidence level (*p* = 0.90) with a 5% margin of error (*r* = 0.05), this translates to 1,082 deaths across the entire experience study (reflecting all ages).

As discussed in Section 3.4.3, experience studies based on amounts-weighting may be more appropriate in setting mortality assumptions for pension valuations. Therefore, instead of using Formula 1, many pension actuaries may wish to determine credibility based on Formula 4 in Section 5.2.2:

Formula 4 (amounts): Full credibility (Z = 1) on an amounts-weighted basis is achieved when

Actual dollars of deaths = $\frac{\lambda}{E_D}\sum_{i=1}^{n}(b_i)^2 q_i^s$

Actual number of deaths = $\frac{\lambda E_N \sum_{i=1}^{n}(b_i)^2 q_i^s}{(E_D)^2}$

Because this equation depends on benefit amounts from a specific plan, the number of deaths required for credibility will vary based on the plan and the benefit cash flows included in any specific experience study.

As discussed previously, there are no formal rules for selecting parameters *r* and *p* in determining full credibility or for using the counts-weighted or amounts-weighted method. Some information can be gleaned from existing standards of practice (as of the date of this publication). For example, the proposed IRS mortality regulations adopt the amounts-weighted approach in applying credibility procedures for constructing plan-specific mortality tables for pension plan valuations. However, this should not be interpreted as an official endorsement of any specific set of assumptions for *all* situations.

The following examples demonstrate how credibility is used to adjust standard mortality tables in cases where there are enough data for full credibility.

Example 1: Data collected for female retirees of Plan A for years 2013–2015 produced 1,617 actual deaths, with an expected number of deaths equal to 1,071. What is the adjustment factor applied to standard female mortality to create the new table, assuming a 95% confidence interval and a 5% margin of error? (Refer to the attached Spreadsheet for detailed calculations—Example 1 (counts).)

Counts-weighted:

Step 1: Based on a normal distribution, *p* = 0.95 translates to a *z*-statistic of 1.96.

Step 2: A 5% margin of error would indicate that full credibility on a counts-weighted basis is achieved if there are $(1.96 / 0.05)^2$ number of deaths. So full credibility would be achieved at 1,537 actual deaths.

Step 3: Since there are 1,617 actual deaths, there are enough for full credibility.

Step 4: Since there is full credibility, Multiple = $\hat{f} = 1617/1{,}071 = 1.51.$

So the new adjusted female table would be 1.51 of the standard table. That is, at each age, ${q_x}^S$ is multiplied by 1.51 to produce ${q_x}^F$. *Note that the standard table reflects the underlying experience study group (females) as well as the midpoint of the study (2014).*

Example 2: Data collected for all retirees of Plan B for years 2013–2015 produced 352 actual deaths, with an expected number of deaths equal to 179. What is the adjustment factor applied to standard mortality to create the new table, assuming a 95% confidence interval and a 5% margin of error? (Refer to the attached Spreadsheet for detailed calculations—Example 2 (amounts).)

Amounts-weighted:

Step 1: Based on a normal distribution, *p* = 0.95 translates to a *z*-statistic of 1.96.

Step 2: A 5% margin of error would indicate that full credibility on an amounts-weighted basis is achieved at 2,352 deaths, based on Formula 4.

Step 3: Since there are 352 deaths, there are not enough for full credibility.

Step 4: See Section 3.4.2, Example 2.

3.4.2 Partial Credibility

From Section 5.2.1:

> **Formula 2:** In the absence of full credibility, *Z* is the square root of the ratio of actual deaths to the number of deaths required for full credibility.

In the case of partial credibility, Multiple is adjusted to shift the final rates closer to the standard table rates based on the equation $Z \cdot (\hat{f}) + (1 - Z) \cdot 1$.

The following examples demonstrate how standard mortality tables are adjusted in the absence of full credibility.

Example 2 (continued): Amounts-weighted (Example 2 (Amounts) in Spreadsheet)

Step 4: Based on Formula 2, $Z = \sqrt{352/2{,}352} = 0.387$.

Step 5: $\hat{f} = A_D/E_D = 4{,}966.2K/3{,}166.1\,K = 1.57$

Step 6: Multiple = 0.387 · 1.57 + (1 – 0.387) · 1 = 1.22

So the new adjusted table would be 1.22 of the standard table. That is, at each age, ${q_x}^S$ is multiplied by 1.22 to produce ${q_x}^F$. *Note that the standard table reflects the underlying experience study group (combined male/female group) as well as the midpoint of the study (2014).*

Example 3: Data collected for male retirees of Plan A for years 2013–2015 produced a total of 971 actual deaths, with an expected number of deaths equal to 1,440. What is the adjustment factor applied to standard mortality rates to create the new table, assuming a 95% confidence interval and a 5% margin of error? (Refer to the attached Spreadsheet for detailed calculations—Example 3 (counts).)

Counts-weighted (Example 3 (Counts) in Spreadsheet)

Step 1: Based on a normal distribution, *p* = 0.95 translates to a *z*-statistic of 1.96.

Step 2: A 5% margin of error would indicate that full credibility on counts-weighted basis is achieved at 1,537 deaths, based on Formula 1.

Step 3: Since there are 971 actual deaths, there are not enough for full credibility.

Step 4: Based on Formula 2, $Z = \sqrt{971/1{,}537} = 0.795$.

Step 5: $\hat{f} = A_N/E_N = 971/1{,}440 = 0.67$

Step 6: Multiple = 0.795 · 0.67 + (1 – 0.795) · 1 = 0.741

So the new adjusted table would be 0.741 of the standard table. That is, at each age, ${q_x}^S$ is multiplied by 0.741 to produce ${q_x}^F$. *Note that the standard table reflects the underlying experience study group (males) as well as the midpoint of the study (2014).*

3.4.3 Amounts- Versus Counts-Weighted

In the LFCT approach, the multiplier is based on the ratio of actual to expected deaths ($\hat{f}$). This ratio can be based on either the number of deaths or the sum of pension amounts of those who have died and are expected to die. The counts-weighted approach requires fewer lives for full credibility and may be easier to apply in practice. However, there are a number of reasons why *using amounts-weighted ratios may be more appropriate for setting the mortality assumption for pension valuations*:

- Pension liabilities are amounts-weighted (i.e., individuals with higher benefit amounts contribute more to the pension liability than those with lower benefit amounts, all else being equal).
- Benefit amounts are often a predictor of mortality rates. Therefore, the estimate will be more accurate to the degree that the distribution of amounts is similar in the future.
- The standard mortality valuation tables (the relevant data that are available) are generally developed using amounts-weighting. So if the experience study does not use amounts-weighting, there may be inconsistencies in the development of the appropriate adjustment.

Consequently, an amounts-weighted actual-to-expected ratio better reflects liability development, may be more accurate and may be more consistent with the relevant data. For example, counts-weighted values may result in a mortality adjustment that leads to understated liabilities.

However, an experience study performed on an amounts basis generally requires more exposures to achieve full credibility than a study based on number of lives. Amounts weighting produces lower credibility because although using amounts may increase the “accuracy” of the liability estimation, it also increases the variance of the estimator. The higher variance leads to less weighting being assigned to the subject experience.

The proposed IRS regulations issued in December 2016 define full credibility using an amounts-weighted approach.

The Appendix shows a detailed development of the LFCT model based on both the counts-weighted and amounts-weighted approaches. In addition, the attached Spreadsheet contains sample formulas for calculating full credibility based on a sample plan’s population.

The number of deaths needed in an amounts-weighted study for full credibility depends on the distribution of the benefit amounts of the participants included in the analysis and is plan-specific. Therefore, generalization of the number of deaths needed for full credibility based on amounts weighting is not possible, unlike in the case of counts-weighting. However, the number of deaths needed for a specific subject population can easily be generated. The following tables show comparisons of numbers of deaths based on counts-weighted and amounts-weighted methodologies, using the sample experience data provided in the attached Spreadsheet.

Number of Deaths Needed for Full Credibility Based on *r* and *p*

	r = 1%		r = 3%		r = 5%	
	Counts	Amounts	Counts	Amounts	Counts	Amounts
p = 90% (z = 1.645)	27,060	41,404	3,007*	4,600	1,082	1,656**
p = 95% (z = 1.96)	38,416	58,788	4,268	6,532	1,537	2,352
p = 99% (z = 2.575)	66,306	101,537	7,367	11,282	2,652	4,061

* Implied full credibility per CIA educational note to insurance practitioners.
** The proposed IRS regulations issued in December 2016 define full credibility using an amounts-weighted approach based on a 90% confidence level and a 5% margin of error.

Number of Deaths Needed for Partial Credibility Assuming *r* = 5% and *p* = 90%

Z	10%*	20%*	30%*	40%	50%	60%	70%	80%	90%	100%
Number of deaths—Counts	11	43	97	173	271	390	530	692	876	1,082
Number of deaths—Amounts	17	66	149	265	414	596	811	1,060	1,341	1,656

* The proposed IRS regulations issued in December 2016 require a minimum of 100 actual deaths in order to apply partial credibility for each customized mortality table.

The statistics for amounts-weighted number of deaths in both of these tables are based on the plan in Example 2 of the attached Spreadsheet. These numbers will differ by plan and experience study years as benefit cash flows differ.

3.5 Actuarial Judgment in Adjusting Standard Table

3.5.1 Is an Adjustment Required?

Before adjusting, the actuary should first evaluate whether it "looks" like an adjustment is even required. One way to evaluate whether an adjustment may be appropriate is to compare actual rates of mortality to the relevant standard table. Graphical analysis is often helpful in seeing whether the table is lower or higher relative to experience. The following example shows two ways of analyzing whether an adjustment may be warranted. In Figure 1, the line Actual to Expected Deaths = 1 indicates that actual is the same as expected (meaning no adjustment is necessary). The red dots indicate the ratio of actual to expected deaths at each age. Figure 1 shows that prior to age 75 (which is where the majority of the plan data are in this example), actual deaths exceed expected deaths, so an upward adjustment may be reasonable. This can also be demonstrated by simply graphing actual mortality experience rates relative to the standard relevant table, as in Figure 2.

Figure 1
Actual to Expected Deaths

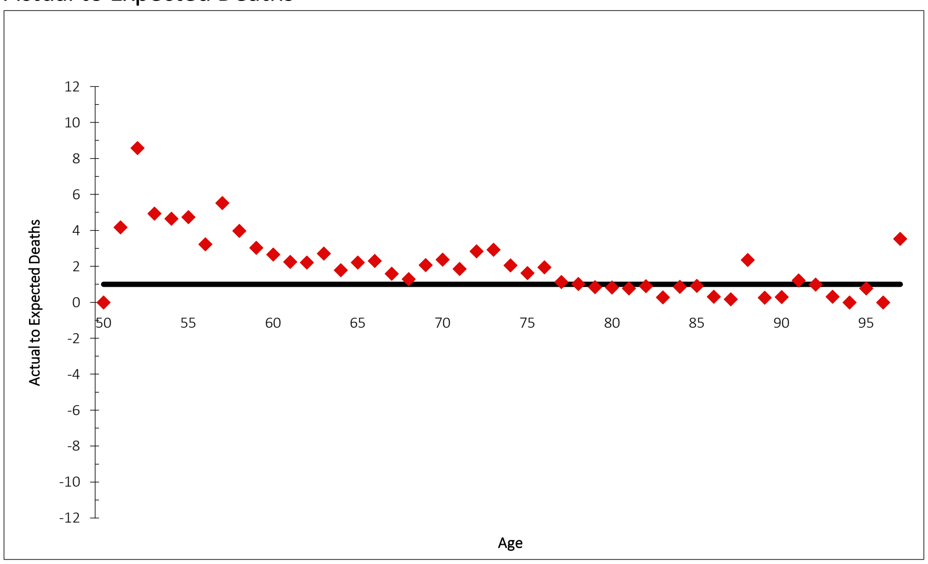
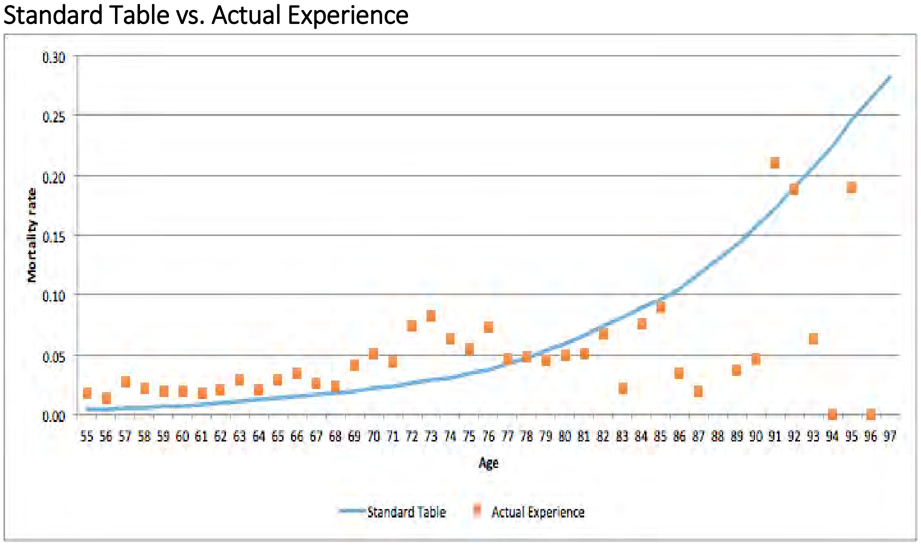


Figure 2
Standard Table vs. Actual Experience



3.5.2 Does the Shape of the Plan's Mortality Experience Match the Shape of the Standard Table?

The preceding figures are also a good way to evaluate whether shifting the whole standard table based on experience is a reasonable approach. As noted earlier, shifting a whole table implies that the shape of the actual mortality table is similar to the shape of the underlying standard table. It is critical to compare the shape at the ages where most of the experience data are collected as well as where most of the plan's liability exists—generally the retirement ages.

In Figure 2, the shape of actual experience prior to age 75, denoted by orange dots, roughly approximates the shape of the underlying standard table. It appears that scaling the table is a reasonable approach in customizing the mortality assumption for this plan.

Methods for addressing the shape of the table are outside the scope of this paper.

3.5.3 Applying Credibility by Age Groupings

The method of adjustment already discussed is based on the assumption that true mortality is equal to a constant multiple of the standard table without any variations. That is, the ultimate shape of the table is the same as the underlying table.

Credibility works best when applied to one parameter; for this purpose, the parameter is a single multiplier of the standard mortality table. It is the overarching assumption in mortality credibility that the same factor applies to the whole table. It is possible to apply the credibility formula to portions of the mortality curve. However, combining the results into a single curve may result in inconsistencies that will likely need to be smoothed using graduation techniques and may be less justifiable from a theoretical standpoint.

It is worth noting that at very old ages (such as above 90), mortality rates do not tend to vary significantly. Consequently, adjusting rates at those ages with an aggregated factor may not be appropriate. Some consideration may be given to limiting adjustments above certain ages. In pension plan valuations, liabilities at older ages tend to be less material, so factoring in extra complexity also may not be worth the effort.

3.5.4 Applying Credibility for Subgroups

As noted earlier, the proposed IRS mortality regulations require that in constructing plan-specific mortality tables, separate mortality credibility analyses be performed for each subgroup within a plan. For example, a plan that contains male, female, white-collar and blue-collar employees will require four different mortality credibility analyses, with each analysis requiring its own separate threshold for full and partial credibility. In this case, each subgroup's own subject experience and standard mortality table will be used to determine the adjustment factor. This type of separate analysis is not required unless a plan is subject to IRS rules. Unless specifically required, in cases where the actuary does not believe mortality rates vary between groups, a combined mortality experience study of all participants may be appropriate. For example, the single threshold of 1,082 deaths across all plan participants could satisfy full credibility requirements, assuming $p = 0.90$ and $r = 0.05$.

3.6 Updating Existing Credibility Analysis

3.6.1 Frequency of Experience Studies

Actuarial judgment is required to assess the frequency of updates to the mortality adjustment. Consideration should be given to the following:

- Significant changes in plan design, population, and/or company-specific shocks since the last review may necessitate a review of more recent mortality experience.
- Trends in industry mortality studies (such as a change in the rate of mortality improvement) may warrant a review of the experience data for companies in the industry.
- The length of time since the last experience study was conducted may be significant. The actuary may look at the midpoint of the last experience data study relative to the current valuation date to see if it is distant enough to require an update. For example, if today's valuation date is 2018, but the last experience study was conducted between 2012 and 2014 with a midpoint of 2013 plan year (five years ago), the actuary might conclude that a new study is appropriate to ascertain the validity of the credibility adjustment.

3.6.2 Selection of a Relevant Experience Base Mortality Table in an Updated Credibility Analysis

When redoing a credibility analysis, it is suggested that the relevant experience table *not* reflect the last credibility adjustment applied to the underlying standard mortality table. The updated analysis should generally incorporate the *most recent and appropriate standard mortality table as of the date of the valuation* because relevant experience is assumed to be "accurate." The credibility adjusted table has some prior subject experience built into it, which decreases its "accuracy" with respect to the new subject experience population. Thus, reflecting outdated plan experience from prior adjustments may skew the current results.

Example 3 (updated experience study 2016–2018): The last study from the original Example 3 produced an adjustment of .741 to the standard valuation table, which was RP-2014 for males (Mort 1). The new study results in actual deaths of 650 (for the period 2016–2018). The new base table is RP-2014 for males with mortality improvement projections through 2017 (Mort 2). This table results in expected deaths of 1,390. The new adjustment factor applied to Mort 2 is equal to 52.2%, as follows:

Step 1: Based on a normal distribution, p = 0.95 translates into a z-statistic of 1.96.

Step 2: A 5% margin of error would indicate that full credibility on a counts-weighted basis is achieved at 1,537 deaths, based on Formula 1.

Step 3: Since there are 650 actual deaths, there are not enough for full credibility.

Step 4: Based on Formula 2, $Z = \sqrt{650/1{,}537}$ =0 .65.

Step 5: $\hat{f} = A_N/E_N = 650/1{,}390 = 0.468$

Step 6: Multiple = 0.468 · 0.65 + (1 – 0.65) · 1 = 0.654

Note: The original adjustment is ignored.

4 Limitations of Adjustments Based on Plan Experience

Adjustments for plan experience are limited by the following considerations:

- Is there enough data to warrant an adjustment based on plan experience? There is no theoretical minimum. Generally, a relatively large data set is needed to factor in subject experience. Since *there is no theoretical basis for minimum plan size*, actuaries should evaluate, from a practical perspective, the size of the credibility factor that would render the impact of actual plan experience so small that it would not be worth analysis.

  The proposed IRS regulations issued in December 2016 require a minimum of 100 actual deaths in order to apply partial credibility in constructing the plan's mortality table based on the plan's experience. Note that the 100-death threshold is applied separately for each group for whom the mortality is being constructed.

- Is the underlying shape of the subject experience different enough from any of the standard tables so that simply shifting the table would not appropriately reflect plan experience? In this case, the actuary should evaluate whether other methods, such as building a table from scratch, should be used, assuming there are sufficient data to consider this.

5 Appendix

5.1 Resources

(Updated as of 12/1/2015)

This list represents a number of resources on credibility theory prepared by SOA staff. These materials provide good background on credibility theory, and several reference other available resources in their bibliographies.

1. Klugman, Stuart, Thom Rhodes, Marianne Purushotham and Stacy Gill. "Credibility Theory Practices Report." Research paper, Society of Actuaries, 2009, http://www.soa.org/research/research-projects/life-insurance/research-credibility-theory-pract.aspx. (This paper contains a bibliography of other resources.)
2. Benjamin, Gavin. "Selecting Mortality Tables: A Credibility Approach." Research paper, Society of Actuaries, 2008, http://www.soa.org/files/research/projects/research-2008-benjamin.pdf.
3. Klugman, Stuart, and Thomas E. Rhodes. "Session 156PD, Application of Credibility Theory." Presented at the SOA 2014 Annual Meeting and Exhibit, Orlando, FL, October 2014, http://www.soa.org/Files/Pd/2014/annual-mtg/2014-orlando-annual-mtg-156-V23.pdf. (The audio for this session is available free through the SOA online store at https://store.soa.org/ProductDetails.aspx?productId=011198. The session has an overview of the revised ASOP and an application to the procedures stipulated for principles-based reserves.)
4. American Academy of Actuaries Pension Committee. "Selecting and Documenting Mortality Assumptions for Pensions." Public policy practice note, revised June 2015, http://www.actuary.org/files/Mortality_PN_060515_0.pdf. (Appendix 2 contains a discussion of credibility theory.)
5. American Academy of Actuaries 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.)
6. Credibility Task Force of the General Committee of the Actuarial Standards Board. "Actuarial Standard of Practice 25: Credibility Procedures." Revised December 2013, http://www.actuarialstandardsboard.org/wp-content/uploads/2014/02/asop025_174.pdf. (Appendix 1 provides a high-level background discussion on credibility practice.)
7. Klugman, Stuart, and Thomas Herzog. "Session 9I: Believe It or Not!" Presented at the SOA Maui II Spring Meeting, Maui, HI, June 1998, http://www.soa.org/library/proceedings/record-of-the-society-of-actuaries/1990-99/1998/january/rsa98v24n29i.aspx. (This session provides a historical look at how credibility was developed during the twentieth century.)
8. The following three sources have been used by the SOA for its preliminary exam covering credibility theory. They all cover the same material:
   - Herzog, Thomas N. *Introduction to Credibility Theory*. 4th ed. New Hartford, CT: Actex Learning/Mad River Books, 2010.
   - Klugman, Stuart A., Harry H. Panjer and Gordon E. Willmot. *Loss Models: From Data to Decisions*. 4th ed. Hoboken, NJ: John Wiley & Sons, 2012.
   - Mahler, Howard C., and Curtis Gary Dean. "Credibility." Chap. 8 in *Foundations of Casualty Actuarial Science*. Casualty Actuarial Society, 2001, http://www.soa.org/files/pdf/C-21-01.pdf.
   - Dean, Curtis Gary. "Topics in Credibility Theory." Study note, 2005, http://www.soa.org/files/pdf/c-24-05.pdf.

5.2 LFCT Full Credibility Model[1]

Assumptions:

- Each life under study is exposed for the full year of age.
- The number of observed deaths (*N*) at age *x* will have a binomial distribution with parameters *N* and *q*, where *q* is the (unknown) probability of death for someone age *x*. For large *N*, binomial distribution can be approximated with a normal distribution.

Notation:

Let *i* denote the *i*th life out of *n* lives studied.

Let q_i^s denote the standard table mortality rate for the *i*th life. Note that the subscript is not the age. The value of *q* is found by looking at that life's age and then applying the value from the standard table.

Let d_i be 0 if the *i*th life survives the year of observation and 1 if the life dies.

Let b_i be the amount associated with the *i*th life.

Then the following quantities can be calculated:

$A_D = \sum_{i=1}^{n} b_i d_i$ = actual death amounts

$A_N = \sum_{i=1}^{n} d_i$ = actual death lives

$E_D = \sum_{i=1}^{n} b_i q_i^s$ = expected death amounts

$E_N = \sum_{i=1}^{n} q_i^s$ = expected death lives

$\hat{f} = A/E$ = observed mortality ratio

Under LFCT, the observed ratio gets full credibility provided it is within *r*% of the true ratio with a probability of at least *p*. That is,

$$\Pr(|\hat{f} - f| \le rf) > p.$$

$$\mu = E(\hat{f}) = \frac{1}{E_D}\sum_{i=1}^{n} f b_i q_i^s = f$$

$$\sigma^2 = Var(\hat{f}) = \frac{1}{E_D^2}\sum_{i=1}^{n} b_i^2 f q_i^s (1 - f q_i^s) \approx \frac{1}{E_D^2}\sum_{i=1}^{n} b_i^2 f q_i^s$$

When we use counts $b_i = 1$, μ and σ^2 are calculated as follows:

$$\mu = E(\hat{f}) = f$$

$$\sigma^2 = Var(\hat{f}) = \frac{1}{E_N^2}\sum_{i=1}^{n} f q_i^s (1 - f q_i^s) \approx \frac{1}{E_N^2}\sum_{i=1}^{n} f q_i^s = \frac{f}{E_N}$$

Note: The approximation is based on the assumption that *q* is very small, so $1 - fq_s$ is close to 1.

---

[1] Based on work by Gavin Benjamin, FSA, FCIA, as reported in "Selecting Mortality Tables: A Credibility Approach" (https://www.soa.org/Files/Research/Projects/research-2008-benjamin.pdf)

5.2.1 Counts-Weighted Derivation

In the counts scenario (i.e., b_i = 1), the requirement for full credibility becomes as follows:

$$\Pr(|\hat{f} - f| \le rf) > p$$
$$\Pr(-rf \le (\hat{f} - f) \le rf) > p$$
$$\Pr\left(-\frac{rf}{\sigma} \le Z \le \frac{rf}{\sigma}\right) > p$$

Note: Here, Z is not the credibility factor but the usual symbol for a standard normal random variable, and z is the percentile of the standard normal distribution, as indicated by the subscript.

$$\frac{rf}{\sigma} > z_{(1+p)/2}$$
$$\frac{rf}{\sqrt{f / E_N}} > z_{(1+p)/2}$$
$$\sqrt{fE_N} > z_{(1+p)/2} / r$$
$$fE_N > (z_{(1+p)/2} / r)^2$$
$$A_N > (z_{(1+p)/2} / r)^2.$$

Note: Because f is unknown, we substitute A_N for fE_N since that is what A_N is estimating.

Formula 1 (counts): Full credibility on a counts-weighted basis is achieved (Z = 1) when

Actual number of deaths = $\lambda = (z_{(1+p)/2}/r)^2$

When A_N is too small, the variance of $\hat{f}$ is too large. Applying the credibility formula (and recalling that for LFCT, the estimate based on relevant experience is not random) reduces the variance: $Var(Z\hat{f} + 1 - Z) = Z^2\sigma^2$. Remember that $\sigma^2 = f/E_N$ and $\sigma = \sqrt{f/E_N}$.

The credibility estimator will have full credibility when

$$\frac{rf}{Z\sigma} = z_{(1+p)/2}$$
$$\frac{rf}{Z\sqrt{f/E_N}} = z_{(1+p)/2}$$
$$\frac{\sqrt{fE_N}}{Z} = z_{(1+p)/2}/r = \sqrt{\lambda}$$
$$\frac{\sqrt{A_N}}{Z} = \sqrt{\lambda}$$
$$Z = \sqrt{A_N/\lambda}$$

Note: Because f is unknown, we substitute A_N for fE_N since that is what A_N is estimating.

Formula 2: In the absence of full credibility, Z is the square root of the ratio of actual deaths to the number of deaths required for full credibility.

A general formula for the credibility factor can be backed out of this development:

Formula 3:

$$Z = \min\left\{1, \frac{r\hat{f}}{z_{(1+p)/2}\hat{\sigma}}\right\}$$

5.2.2 Amounts-Weighted Derivation

There is no simplification this time due to the *b*-squared terms. Using Formula 3, the credibility factor is

$$Z = \min\left\{1, \frac{r\hat{f}}{z_{(1+p)/2}\hat{\sigma}}\right\} = \min\left\{1, \frac{r}{z_{(1+p)/2}} \frac{f}{\sqrt{\sum_{i=1}^{n} b_i^2 f q_i^s / E_D^2}}\right\} = \min\left\{1, \frac{r}{z_{(1+p)/2}} \frac{\sqrt{f} E_D}{\sqrt{\sum_{i=1}^{n} b_i^2 q_i^s}}\right\}.$$

From this, we can infer the standard for full credibility. Letting $\lambda = (z_{(1+p)/2} / r)^2$, there is full credibility when

$$\frac{1}{\sqrt{\lambda}} \frac{\sqrt{f} E_D}{\sqrt{\sum_{i=1}^{n} b_i^2 q_i^s}} \geq 1$$

$$\frac{f E_D^2}{\sum_{i=1}^{n} b_i^2 q_i^s} \geq \lambda$$

$$A_D \geq \frac{\lambda}{E_D} \sum_{i=1}^{n} b_i^2 q_i^s.$$

$$A_N = \frac{A_D}{\frac{E_D}{E_N}} \geq \frac{\lambda E_N \sum_{i=1}^{n} b_i^2 q_i^s}{E_D^2}$$

Formula 4 (amounts): Full credibility on amounts-weighted basis is achieved (Z = 1) when

Actual dollars of deaths = $\frac{\lambda}{E_D} \sum_{i=1}^{n} (b_i)^2 q_i^s$

Actual number of deaths = $\frac{\lambda E_N \sum_{i=1}^{n} (b_i)^2 q_i^s}{(E_D)^2}$

5.3 Overview of the Bühlmann Credibility Model

This model is based on the following principles:

1. The estimate for the mean takes the form $\hat{m} = Z\overline{x} + (1-Z)\hat{\mu}$, where $\overline{x}$ is the mean of the subject data ($x_1, x_2, \ldots, x_n$), and $\hat{\mu}$ is an estimate of the mean of the relevant data.
2. The two estimates ($\overline{x}$ and $\hat{\mu}$) are both uncertain. The error in the first estimate is termed the process variance, and the error in the second is the variance in the hypothetical means.
3. The optimal solution provides more credibility when the process variance is smaller or the sample size is larger (both of which imply that the sample mean is more accurate) and when the variance of the hypothetical means is larger (which indicates that the relevant experience is less reliable).
4. The result is that $Z = n/(n+K)$, where *n* is the number of observations in the subject-specific data set, and *K* = process variance/variance of the hypothetical means.
5. One way to obtain this result is to minimize the expected squared error of the estimate where the expectation is taken over all possible observed values as well as all possible underlying parameters for the subject experience.

To use this model, the following values need to be determined:

- An estimate of the population mean over all possible groups of subject experience
- An estimate of the expected process variance, which is the variance of a single observation from a single entity (plan) averaged over all such plans
- An estimate of the variance of hypothetical means, which is the variability of average experience over all subject groups

More detailed information on the Bühlmann model, including proofs and examples of the approach, can be found in these resources:

- Herzog, Thomas N. *Introduction to Credibility Theory*. 4th ed. New Hartford, CT: Actex Learning/Mad River Books, 2010.
- Klugman, Stuart A., Harry H. Panjer and Gordon E. Willmot. *Loss Models: From Data to Decisions*. 4th ed. Hoboken, NJ: John Wiley & Sons, 2012. (Four editions are available and cover credibility with only slight variations.)
- Mahler, Howard C., and Curtis Gary Dean. "Credibility." Chap. 8 in *Foundations of Casualty Actuarial Science*. Casualty Actuarial Society, 2001, http://www.soa.org/files/pdf/C-21-01.pdf.
- Dean, Curtis Gary. "Topics in Credibility Theory." Study note, 2005, http://www.soa.org/files/pdf/c-24-05.pdf.

5.4 Mortality Tables

RP-2014 Male Female Employee Healthy annuitant Disabled	**RP-2000** Male Female Employee Healthy annuitant Combined healthy Disabled	**CPM 2014** Male Female Private Public
RP-2014 (Blue Collar) Male Female Employee Healthy annuitant	**RP-2000 (Blue Collar)** Male Female Combined healthy	**Older Tables** GAM 71 GAM 83 UP 84* UP 94* GAR 94 GAM 94
RP-2014 (White Collar) Male Female Employee Healthy annuitant	**RP-2000 (White Collar)** Male Female Combined healthy	
RP-2014 (Bottom Quartile) Male Female Employee Healthy annuitant	**RP-2000 (Small Benefit Amount)** Male Female Healthy annuitant	
RP 2014 (Top Quartile) Male Female Employee Healthy annuitant	**RP 2000 (Medium Benefit Amount)** Male Female Healthy annuitant	
RP 2014 (Juvenile) Male Female	**RP 2000 (Large Benefit Amount)** Male Female Healthy annuitant	

* Unisex tables are a 50-50 combination of male and female.

About The Society of Actuaries

The Society of Actuaries (SOA), formed in 1949, is one of the largest actuarial professional organizations in the world dedicated to serving 28,000 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 policy makers 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.

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