1 Actuarial Science

Stuart Klugman

1.1 Introduction

Actuarial science is an applied mathematical and statistical discipline in which data-driven models are constructed to quantify and manage financial risk. The term "actuarial statistics" is not in common use because a well-defined set of statistical techniques useful to actuaries has not been established. This topic could also be viewed as a discussion of the types of data (mortality, morbidity, accident frequency, and severity) collected by actuaries, but that will not be the focus here. This entry will concentrate on two particular statistical endeavors in which actuaries have played a major role—construction of mortality tables and credibility theory.

1.2 Construction of Mortality Tables

From the 1600s, governments sold annuities based on individuals' lifetimes. To be useful as a fundraising mechanism, the cost of the annuity needed to be greater than the expected cost of the benefit. Although not the first mortality table (or life table), the work of Halley [10] combined the construction of a mortality table with the concept of expected present value. From the life table, for a person of current age x, it is possible to get the probability distribution of the number of years remaining; that is,

$$ {}_{k|}q_x = \Pr(\text{death is between ages } x+k \text{ and } x+k+1),\ k=0,1,\ldots. $$

In addition, if the annuity is to pay one monetary unit at the end of each year, provided the annuitant is alive, the expected present value is

$$ a_x = \sum_{k=1}^{\infty} {}_{k|}q_x (v + v^2 + \cdots + v^k), $$

where $v = 1/(1+i)$ and i is the rate of interest.

A few years later, de Moivre [19] introduced an approximation based on linear interpolation between values in the life table (his table did not have survival probabilities at each integral age). This approximation continues to be used today and is referred to as the *uniform distribution of deaths* assumption [4, Chap. 3].

Life tables for actuarial use were constructed on an adhoc basis until the middle of the twentieth century when the so-called "actuarial method" was developed. It is loosely based on an assumption put forth by Balducci [2], viz.,

$$ \Pr(\text{a person age } x+t \text{ dies before age } x+1) = (1-t)\Pr(\text{a person age } x \text{ dies before age } x+1), $$

$0 < t < 1$. The result is an exposure-based formula that estimates the key life-table quantity as

$$\begin{aligned} q_x &= \Pr(\text{a person age } \times \text{ dies before age } x+1) \\ &= \text{number of observed deaths/exposure.} \end{aligned}$$

For a life observed between ages x and $x+1$, the exposure contribution is the portion of the year the life was observed, except for deaths, for which the exposure is the time from first observation to age $x+1$.

This estimator is inconsistent [5]. However, it has one quality that made it extremely valuable. Given the types of records commonly kept by insurance companies, this formula was easy to implement by hand, or using mainframe computers prevalent through the 1980s. A good exposition of the actuarial method and its practical applications is Reference 3. Since then, actuaries have used the more accurate Kaplan-Meier [13] and maximum likelihood estimation* procedures. These concepts are introduced in an actuarial setting in Reference 6.

Once the values of q_x have been obtained, a second actuarial contribution has been the smoothing of these values to conform with the *a priori* notion that from about age five onward the values should be smoothly increasing. The process of smoothing mortality rate estimates is called *graduation*. An introduction to all of the commonly used methods is given in Reference 17. Two of the more commonly used methods, interpolation, and Whittaker, will be discussed here. Both methods create the graduated rates as a linear combination of surrounding values.

The interpolation method requires that the observations be grouped in a manner that creates estimates of q_x at every k (often 5) years of age. This is done by first aggregating the deaths (say, d_x) and exposures (say, e_x) at the surrounding ages, to create, for example, with $k = 5$,

$$\begin{aligned} d_x^* &= d_{x-2} + d_{x-1} + d_x + d_{x+1} + d_{x+2}, \\ e_x^* &= e_{x-2} + e_{x-1} + e_x + e_{x+1} + e_{x+2}. \end{aligned}$$

Because these series are often convex, an improved aggregated value can be found from King's pivotal point formula [14]:

$$\begin{aligned} d_x^{**} &= -0.008 d_{x-5}^* + 0.216 d_x^* \\ &\quad -0.008 d_{x+5}^*, \\ e_x^{**} &= -0.008 e_{x-5}^* + 0.216 e_x^* \\ &\quad -0.008 e_{x+5}^*. \end{aligned}$$

Finally, the mortality rate at age x is given by $q_x^{**} = d_x^{**}/u_x^{**}$.

The most commonly used interpolation formula is the Karup-King formula

$$\begin{aligned} & q_{x+j} \\ &= s q_{x+5}^{**} + 0.5 s^2 (s-1) \delta^2 q_{x+5}^{**} \\ &\quad + (1-s) q_x^{**} + 0.5 (1-s)^2 (-s) \delta^2 q_x^{**}, \\ &\qquad j = 0, 1, 2, 3, 4, 5, \end{aligned}$$

where $s = j/5$ and $\delta^2 q_x^{**} = q_{x+5}^{**} - 2q_x^{**} + q_{x-5}^{**}$ is the second central difference. This formula uses four mortality rates and has the property that if those rates lie on a quadratic curve, the interpolated values will reproduce that curve. In addition, the cubic curves that connect consecutive mortality rates will have identical first derivatives where they meet.

Another popular formula is due to Jenkins [12]. It requires fourth central differences and thus involves six points. It reproduces third-degree polynomials and adjacent curves will have identical first and second derivatives. To achieve these goals, the formula does not match the original mortality rates. That is, $q_{x+0} \neq q_x^{**}$.

The Whittaker method [22] can be derived by a Bayesian argument or from arguments similar to those used in creating smoothing splines. Let q_x, $x = 0, \ldots, n$, be the original estimates; let v_x, $x = 0, \ldots, n$, be the graduated values; and let w_x, $x =$

$0, \ldots, n$, be a series of weights. Then, the graduated values are those that minimize the expression

$$\sum_{x=0}^{n} w_x(v_x - q_x)^2 + h \sum_{x=0}^{n-z} (\Delta^z v_x)^2.$$

The weights are often chosen as either the exposure (sample size) at each age or the exposure divided by $q_x(1 - q_x)$, which would approximate using the reciprocal of the variance as the weight. The value of z controls the type of smoothing to be effected. For example, $z = 3$ leads to graduated values that tend to follow a quadratic curve. The choice of h controls the balance between fit (having the graduated values be close to the original values) and smoothing (having the graduated values follow a polynomial).

1.3 Credibility

Credibility theory is used by actuaries in the setting of premiums based on prior or corollary information. Two common situations are *experience rating* and *classification ratemaking*. An example of the former is workers compensation insurance. Suppose a particular employer had been charged a standard rate on the basis of expecting $\$x$ of claim payments per thousand dollars of payroll. In the previous year, the employer had claims of $\$y$ per thousand dollars of payroll, where $y < x$. The employer believes that a reduction in premium is warranted, while the insurer may claim that the result was simply good fortune. A credibility procedure will base the next premium on the value $zy + (1 - z)x$, where $0 \leq z \leq 1$ and z is called the *credibility factor*. The magnitude of z is likely to depend on the sample size that produced y, the variance of y, and, perhaps, some measure of the accuracy of x.

With regard to classification ratemaking, consider setting premiums for automobile insurance. Separate rates may be needed for various combinations of gender, age, location, and accident history. Let y be an estimate based on the data for a particular combination of factors, and let x be an estimate based on all the data. Because some combinations may occur infrequently, the reliability of y may be low. A credibility estimate using $zy + (1 - z)x$ may be more accurate (although biased). Credibility analysis succeeds for just that reason. By applying the factor $1 - z$ to an estimator that is more stable, the reduction in variance may offset the effect of bias, producing a smaller mean square error.

Two approaches to credibility have evolved. One, usually credited to Mowbray [20], has been termed *limited fluctuation credibility*. The question reduces to determining the sample size needed so that the relative error when estimating the mean will be less than $k\%$ with probability at least $p\%$. A normal or Poisson approximation along with a variance estimate is usually sufficient to produce the answer. If the sample size exceeds this number, then $z = 1$ (full credibility) is used. If not, z is customarily set equal to the square root of the ratio of the actual sample size to that needed for full credibility. Assuming no error in the quantity being multiplied by $1 - z$, the effect is to reduce the variance to equal that which would have been obtained with the full credibility sample size. The simplicity of this method causes it to remain popular. Its drawback is that it does not allow for the increased bias as z decreases, nor does it allow for any error in the quantity being multiplied by $1 - z$.

The second method has been termed *greatest accuracy credibility* and bears a strong resemblance to Bayesian analysis. It was introduced by Whitney [21] with a more thorough derivation produced by Bailey [1] and a modern derivation by Bühlmann [7]. This approach begins by assuming that a sample of size n is obtained from an individual. The observations are independent realizations of the random variable X with a distribution function that depends on the vector parameter θ. Define

$$E(X|\theta) = \mu(\theta) \text{ and } \mathrm{Var}(X|\theta) = v(\theta).$$

Furthermore, assume that θ is unknown, but has been drawn at random from a random variable Θ with distribution function $F_\Theta(\theta)$. Finally, assume that $\mu(\theta)$ is to be estimated by a linear function of the observations; that is,

$$\widehat{\mu(\theta)} = \alpha_0 + \alpha_1 X_1 + \cdots + \alpha_n X_n.$$

The objective is to minimize

$$E_{\Theta, X_1, \ldots, X_n}\left\{\left[\widehat{\mu(\Theta)} - \mu(\Theta)\right]^2\right\}.$$

That is, the squared error should be minimized both over all possible observations and all possible parameter values. For a particular insured with a particular value of θ, the squared error may be larger than if the sample mean were used, but for others, it will be smaller so that the overall error is reduced.

The solution is

$$\widehat{\mu(\theta)} = z\bar{x} + (1-z)\mu;\ \mu = E[\mu(\Theta)],$$
$$z = \frac{n}{n+k},\ k = \frac{E[v(\Theta)]}{\mathrm{Var}[\mu(\Theta)]}.$$

It turns out to be the Bayesian (posterior mean) solution for certain common cases such as normal-normal and Poisson-gamma.

In practice, the indicated quantities must usually be estimated. An approach given in Bühlmann and Straub [8] provides an empirical Bayes estimate, derived by a method of moments approach. This is not unreasonable because the distribution of Θ is not an *a priori* opinion, but rather a real, if unobservable, distribution of how characteristics vary from policyholder to policyholder or group to group. With data on several policyholders or groups, it is possible to estimate the needed moments. A true Bayesian model can be constructed by placing a prior distribution on the parameters of the distribution of Θ. This is done for the normal-normal model in Reference 15.

Textbooks that develop these credibility topics and more (all include an English language version of the Bühlmann-Straub formula) include References 9, 11, 16, and 18 as well as Chapter 5. A comprehensive list of book and article abstracts through 1982 is found in Reference 23.

References

1. Bailey, A. (1950). Credibility procedures. *Proc. Casualty Actuarial Soc.*, **37**, 7–23, 94–115.
2. Balducci, G. (1921). Correspondence. *J. Inst. Actuaries*, **52**, 184.
3. Batten, R. (1978). *Mortality Table Construction*. Prentice-Hall, Englewood Cliffs, NJ.
4. Bowers, N., Gerber, H., Hickman, J., Jones, D., and Nesbitt, C. (1997). *Actuarial Mathematics*, 2nd ed. Society of Actuaries, Schaumburg, IL.
5. Breslow, N. and Crowley, J. (1974). A large sample study of the life table and product limit estimates under random censorship. *Ann. Stat.*, **2**, 437–453.
6. Broffitt, J. (1984). Maximum likelihood alternatives to actuarial estimators of mortality rates. *Trans. Soc. Actuaries*, **36**, 77–142.
7. Bühlmann, H. (1967). Experience rating and credibility. *ASTIN Bull.*, **4**, 199–207.
8. Bühlmann, H. and Straub, E. (1970). Glaubwürdigkeit für Schadensätze (credibility for loss ratios). *Mitteilungen der Vereinigung Schweizerisher Versicherungs-Mathematiker*, **70**, 111–133. (English translation (1972) in *Actuarial Research Clearing House*.)

9. Dannenburg, D., Kass, R., and Goovaerts, M. (1996). *Practical Actuarial Credibility Models.* Ceuterick, Leuven, Belgium.
10. Halley, E. (1694). An estimate of the degrees of the mortality of mankind, drawn from curious tables of births and funerals at the city of Breslau; with an attempt to ascertain the price of annuities on lives. *Philos. Trans.*, **17**, 596–610.
11. Herzog, T. (1996). *Introduction to Credibility Theory.* Actex, Winsted, CT.
12. Jenkins, W. (1927). Graduation based on a modification of osculatory interpolation. *Trans. Am. Soc. Actuaries*, **28**, 198–215.
13. Kaplan, E. and Meier, P. (1958). Nonparametric estimation from incomplete observations. *J. Am. Stat. Assoc.*, **53**, 457–481.
14. King, G. (1887). Discussion: The graphic method of adjusting mortality tables (by T. Sprague). *J. Inst. Actuaries*, **26**, 114.
15. Klugman, S. (1987). Credibility for classification ratemaking via the hierarchical linear model. *Proc. Casualty Actuarial Soc.*, **74**, 272–321.
16. Klugman, S., Panjer, H., and Willmot, G. (1998). *Loss Models: From Data to Decisions.* Wiley, New York.
17. London, D. (1985). *Graduation: The Revision of Estimates.* Actex, Winsted, CT.
18. Mahler, H. and Dean, C. (2001). *Credibility.* In *Foundations of Casualty Actuarial Science*, 4th ed. Casualty Actuarial Society, Arlington, VA.
19. de Moivre, A. (1725). *Annuities Upon Lives.* Fayram, Motte and Pearson, London.
20. Mowbray, A. (1914). How extensive a payroll exposure is necessary to give a dependable pure premium? *Proc. Casualty Actuarial Soc.*, **1**, 24–30.
21. Whitney, A. (1918). The theory of experience rating. *Proc. Casualty Actuarial Soc.*, **4**, 274–292.
22. Whittaker, E. and Robinson, G. (1924). *The Calculus of Observations.* Blackie and Sons, London.
23. de Wit, G., ed. (1986). Special issue on credibility theory. *Insurance Abstr. Rev.*, **2**(3).