Some Consequences of Mortality Function

Farrokh Guiahi

1. Introduction

Actuaries are often interested to find out whether the form (level) of mortality function of their interest has changed significantly. Another issue of importance to the actuaries is determining margins for important Valuation Tables. In this paper, by a special emphasis on the notion of mortality function, procedures are to be considered which should enhance our understanding with regard to the above problems.

Our main shortcoming is that no simple procedure (or algorithm) is provided. However initial stages are set for further research on this important topic.

2. Notations and preliminaries

Actuaries are often concerned with the values of q_x, s, that is the probability of dying within a year after attaining age x. We define the mortality function q as

$$q: \ [0,\infty) \longrightarrow [0,1]$$

such that the value of q for the age x is equal to q_x. Let us denote by q^o our initial (standard) mortality function which we assume to be fairly smooth. Let Y denote the data used to derive the observed mortality function q^1. In order to emphasize the dependence of q^1 on Y we shall write $q^1(Y)$ and $q_x^1(Y)$ is the observed mortality rate for age x. The symbol Y which signifies the data can be quite complex containing information with regard to observed deaths and exposures. It is assumed that such a Y is stochastically generated. That is, if the same group of persons are observed under identical conditions, as the experience which generated Y, the value associated with the new realization of Y would not necessarily be the same. Thus $q^1(Y)$ is a <u>random function</u> and its value for a given x, $q_x^1(Y)$, is a <u>random variable.</u> Furthermore, we assume that q^o, our initial mortality function, which explains the Law of Mortality is non-stochastic.

3. Distance function

In order to detect a change in the level of the mortality function we must examine to see whether q^o and $q^1(Y)$ are close. This necessitates the introduction of a distance function, as we are comparing the closeness of the two functions q^o and $q^1(Y)$. Let Q denote the set of mortality functions for a given class of insurance business or for a specified group of persons.

We define d, a distance function as follows:

$$d: Q \times Q \longrightarrow R,$$ where R denotes the set of real numbers.

d has the following properties:

(i) Non-negativity. For any q, $q'' \in Q$

$d(q', q'') \geqslant o$, and $d(q', q'') = o$ if and only if $q' = q''$.

(ii) Symmetry. For any $q', q'' \in Q$

$d(q', q'') = d(q'', q')$.

(iii) Triangle inequality. For any $q', q'', q''' \in Q$

$d(q', q'') \leqslant d(q', q''') + d(q''', q'')$.

The pair (Q,d) defines a metric space, Royden (1970).

Now for illustrative purposes let us define some distance functions of interest.

(a) $d_1(q', q'') = \left[\sum (q'_x - q''_x)^2\right]^{\frac{1}{2}}$

(b) $d_2(q', q'') = \sum \left| q'_x - q''_x \right|$

(c) $d_3(q', q'') = \max_x \left| q'_x - q''_x \right|$

We note the use of d_1 in the Whittaker - Henderson Type A formulas as a measure of Fit, Miller (1946).

Since the observed mortality function, q^1 is a random function then $d(q^1(Y), q^o)$ is a random variable.

The actuary may be interested in the following probability:

$$P(d(q^1(Y), q^o) \leqslant c_\alpha) = 1-\alpha$$

where α is a significance level, say o.10. To evaluate the above probability we need to know how Y is stochastically generated and also the form of the distance function d. If we are considering a Statistical Hypothesis Test $H_o : q = q^o$ against $H_1 : q \neq q^o$, then we may view q^o as representing the "true level of mortality function".

Given a value of α and <u>if</u> the value of c_α is known then the actuary would compute $d(q^1(Y), q^o)$ and if this value is in excess of c_α the conclusion is reached that the level (form) of the mortality function has significantly changed.

4. Confidence regions

Figure 1, below, will greatly facilitate our understanding with regard to the "confidence regions".

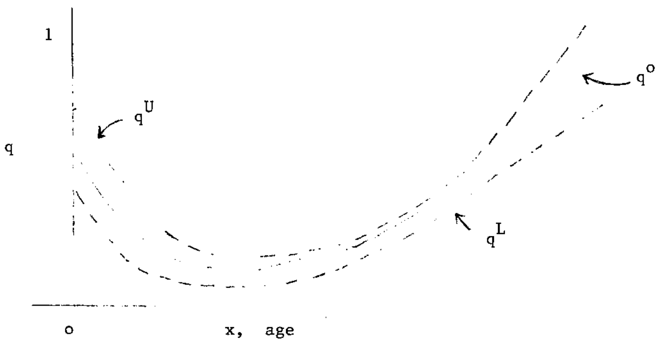

Figure 1.

Figure 1 illustrates the plot of three mortality functions namely q^o (our initial mortality function), q^L, and q^U. L and U are used for the Lower and Upper mortality functions.

Define a set S(subset of Q, set of all mortality functions of interest) as follows:

$$S= \left\{ q \in Q: q_x^L \leqslant q_x \leqslant q_x^U \quad \text{for all } x \geqslant o \right\}$$

From the above diagram, any q which belongs to S lies between the lower and upper mortality functions. In particular $q^o \in S$. It is also possible that q^L and q^U are stochastically determined.

The actuary may then be interested in the following probability

$$P(q^1(Y) \in S)= 1- \alpha$$

where as before α is the significance level and q^1 is the observed mortality function.

The curves q^L and q^U are chosen in order to protect against possible fluctuations inherent in the data. Thus, their selections is directly related to the problem of setting suitable margins for q_x,s.

5. Immediate application possibilities

A special case of the preceding analysis is now considered. Let $x_1, x_2, \ldots, x_k$ denote ages of most interest to the actuary in his mortality investigations. Let $[a_i,b_i]$ denote the closed interval between points a_i and b_i $(b_i > a_i)$; and let us write $q_i^1(Y)$ instead of $q_{x_i}^1(Y)$. Then, $q_1^1(Y), \ldots, q_k^1(Y)$ are k random variables.

In considering the problem of the margins (fluctuations) we may consider the following probability $P(q_i^1(Y) \in [a_i,b_i]$, $i=1,2,\ldots,k)$.
Such a probability should have a large value (close to one) in order to protect against differing experiences. Also note that the above probability is a statement about a collection of random variables, i.e. $q_1^1(Y),\ldots,q_k^1(Y)$. These random variables are not necessarily independent.

Let us choose the a_i,s and b_i,s in the following fashion:

$$a_i = \max \left\{ q_i^o - \alpha_i \cdot \sigma(q_i^1),\ o \right\},$$

$$b_i = \min \left\{ q_i^o + \beta_i \cdot \sigma(q_i^1),\ 1 \right\},$$

where α_i and β_i, $i = 1,2,\ldots,k$, are chosen variables and $\sigma(q_i^1)$ is the standard deviation of the random variable $q_i^1(Y)$. The values of α_i and β_i's are not necessarily equal for each i and may very considerably for different ages (X_i 's).

Define $c_i = P(q_i^1(Y) \notin [a_i,b_i]$), that is c_i is the probability that $q_i^1(Y)$ does not belong to the interval $[a_i,b_i]$. By using the Bonferron's inequality, Bickel & Doksum (1977), we have that

$$P(q_i^1(Y) \in [a_i,b_i] \quad, i=1,2,\ldots,k) \geqslant 1 - \sum_1^k c_i$$

Under certain assumptions, $q_i^1(Y)$ has a binomial distribution, Batten (1978), and its variance is

$$\sigma^2(q_i^1) = \left[q_i^o \ (1-q_i^o) \right] / E_x,$$

where E_x is the exposure. Hence, $\sigma(q_i^1)$ may be approximated by

$$\left[q_i^1(Y) \cdot (1-q_i^1(Y))/E_x \right]^{\frac{1}{2}}$$

By selecting α_i and β_i 's, for example $\alpha_i=\beta_i = 2$, then c_i, s can be computed (also by the use of normal approximation). Thus a lower bound is found for the above probability.

The actuary is in a position to specify α_i and β_i 's according to his concern with regard to possible fluctuations.

6. Conclusion

Further analytic results on these topics are welcome and the possibility of computing some of the above probabilities by the method of Monte Carlo should be further explored.

Acknowledgement.

I am grateful to the many academic actuaries who commented on an earlier draft of this paper. In particular I would like to thank Professors James C. Hickman and Donald A. Jones. I regret that I have not been able to incorporate all their suggestions in this revised version.

References

Batten, R.W. (1978). Mortality Table Construction. Prentice-Hall Inc.

Bickel, P.J. & Doksum, K.A. (1977). Mathematical Statistics: Basic Ideas and Selected Topics. Holden-Day, Inc.

Miller, M.D. (1946). Elements of Graduation. Society of Actuaries (Chicage)

Royden, H.L. (1970). Real Analysis. The Macmillan Company.

Farrokh Guiahi
P.O. Box 11-1859
Tehran, Iran
Revised July 1979