CAMBRIDGE

Cambridge University Press
978-1-316-60191-4 - Some Recent Researches in the Theory of Statistics and Actuarial Science
J. F. Steffensen
Excerpt
More information

First Lecture

1. When, some time ago, I had the honour of receiving an invitation to deliver in the University of London a course of lectures on a subject connected with Statistics, I felt a certain hesitation about accepting, because I am not so much a statistician as an actuary. As, however, the subject on which I was invited to lecture was not termed "statistics" without further qualification, but was only to be a subject "connected with" statistics, I thought I might accept after all, as most of what I have written is concerned with that borderland between statistics and mathematics which constitutes actuarial science. I therefore propose to give an account of some of the efforts I have made to introduce more rigour into certain questions of theoretical statistics and actuarial science, drawing my examples from widely different sources. It will be convenient to begin with a few remarks about the place of mathematics in statistical and actuarial science.

The first point of view that occurs to the mind is that mathematics, even when applied to observed data, is a science that investigates *the relations which exist between numbers*. Observations may contradict each other, owing to unavoidable errors of observation, but mathematical relations are not allowed to contain contradictions. Statistical and actuarial theory must therefore always be presented in such a form that the theoretical relations or assumptions contain no contradiction. In this first lecture I intend to show by examples how we may be led astray by neglect of this principle.

In the second place, mathematics is the proper instrument for justifying *methods of numerical approximation*. Such methods frequently originate in practical work where some approximate method has been found to produce satisfactory results. But too often the computer leaves the matter at that and takes it for granted that the results will be equally satisfactory in other cases. He treats the problem as a statistical one, while its nature is purely mathematical. More or less consciously he obliterates the profound difference between interpolation and graduation, and combines both into a single calculus of observations. It is hardly an exaggeration to say that it is a universal habit amongst actuaries and statisticians to regard a formula of approximation as definitely established when good results have been obtained in a few trial cases. In the second lecture we will consider some questions of this nature, and also occupy ourselves with the allied subject of numerical inequalities.

Thirdly, mathematics is employed for *describing facts of observation*. The formulas used in statistics for this purpose are often of an entirely empirical nature. But there are also cases where theoretical reasons can be given for believing that one formula will fit the facts better than another; and much work can then be saved by choosing from the outset the most suitable formula. It is therefore of considerable practical importance to investigate the theoretical foundation of formulas derived by speculation. Striking examples of such formulas are the types of frequency-functions which will be discussed in the third and last lecture.

2. I shall begin by a critical examination of the notion of *Biometric Functions*. I have dealt with this subject on an earlier occasion*, but before a mathematical rather than a statistical audience, so that I feel justified in recapitulating my views here and illustrating them with an application.

Before going into the objections which can be raised against the manner in which these questions are usually dealt with, I will rather introduce the biometric functions in a way which does not seem open to serious objection.

Let us consider a group of individuals of the same age x and selected according to the same principle with respect to the other essential factors affecting mortality, so that the group may be looked upon as an aggregate of *repeated observations*. Under these circumstances we may assume the existence of such a function μ_x, continuous for all ages $x > 0$, that $\mu_x dx$ represents the probability for a life aged x of dying between the ages x and $x + dx$. The existence of this function—the *force of mortality*—is a postulate, but one which is supported by the evidence of experience; for the causes of death are either constant for all ages (many forms of accidents), or else dependent on the way in which the organism develops and finally wears out, and this process is of a continuous nature.

Very little can be said *a priori* about the function μ_x. Perhaps the only statement that can be made without consulting mortality observations is that there must exist a positive constant ϵ, *independent of* x, such that

$$\mu_x \geqslant \epsilon > 0 \qquad (x \geqslant 0). \qquad \ldots\ldots(1)$$

For a lower limit to μ_x, greater than zero, can at least be derived from the probability of dying by accident. The simple fact expressed by (1) is, however, as we shall presently see, of considerable importance.

* *Proceedings of the Sixth Scandinavian Congress of Mathematicians*, pp. 329–343.

By means of the force of mortality the other biometric functions may be obtained as follows. Let there be l_x persons alive at age x. The mathematical expectation of death amongst these persons in the interval from x to $x + dx$ is $l_x \mu_x dx$. The expected number of living at the end of the interval dx is, therefore,

$$l_{x+dx} = l_x - l_x \mu_x dx,$$

whence $$\mu_x = -\frac{1}{l_x}\frac{dl_x}{dx} = -D \operatorname{Log} l_x, \qquad \text{......(2)}$$

where D denotes the operation of differentiation, and Log the natural logarithm.

Integrating (2), we obtain for the probability that a person aged x is alive after the time t

$$_tp_x = \frac{l_{x+t}}{l_x} = e^{-\int_x^{x+t} \mu_x dx}, \qquad \text{......(3)}$$

whence in particular, for $t = 1$,

$$p_x = e^{-\int_x^{x+1} \mu_x dx} \qquad \text{......(4)}$$

while the probability of dying within a year is $q_x = 1 - p_x$.

If, in (3), we put $x = \alpha$ and thereafter $t = x - \alpha$, we find

$$l_x = l_\alpha e^{-\int_\alpha^x \mu_x dx}. \qquad \text{......(5)}$$

3. Before proceeding, let us see what general conclusions can be drawn from these results, concerning the nature of the biometric functions.

It follows from (3) and (1) that, as t increases, the probability of surviving, $_tp_x$, decreases in a monotonic sense to zero without attaining this value for any finite value of t.

From (1) and (5)—where l_α may be considered as an arbitrary constant—it may be concluded that, as x increases, the function l_x decreases in a monotonic sense to zero without attaining this value for any finite value of x. We may even say something about the rapidity of this decrease; for we have

$$l_x \leqslant l_\alpha e^{-(x-\alpha)\epsilon} \qquad (x \geqslant \alpha), \qquad \text{......(6)}$$

which shows that the decrease is, at least, so rapid that all the *moments*

$$\int_\alpha^\infty x^r l_x dx$$

and the *repeated integrals*

$$\int_\alpha^\infty \int_x^\infty \cdots \int_x^\infty l_x dx^r$$

are necessarily convergent.

The interesting question, whether $\mu_x \to \infty$ as $x \to \infty$, cannot be decided either by observation or by speculation. We obtain from (4), by the theorem of mean value,

$$q_x = 1 - e^{-\mu_\xi} \quad (x < \xi < x + 1). \qquad \ldots\ldots(7)$$

From this relation, it follows that $\mu_x \to \infty$ as $q_x \to 1$ and *vice versa*. But whether $q_x \to 1$ as $x \to \infty$ is impossible to decide. All that can be said is that this assumption, if desired, can be made without introducing any contradiction.

There is every reason to believe that above a certain age the function μ_x *does not decrease*. As, by (4),

$$-\operatorname{Log} p_x = \int_x^{x+1} \mu_x dx,$$

we have under these circumstances

$$\mu_x \leqslant -\operatorname{Log} p_x \leqslant \mu_{x+1}. \qquad \ldots\ldots(8)$$

But, as

$$-\operatorname{Log} p_x = -\operatorname{Log}(1 - q_x) = q_x + \tfrac{1}{2} q_x^2 + \tfrac{1}{3} q_x^3 + \ldots,$$

it follows from (8) that

$$q_x < \mu_{x+1}, \qquad \ldots\ldots(9)$$

provided only that μ_x does not decrease in the interval from x to $x + 1$. As q_x and μ_x do not differ greatly for the ages which are of practical importance, this simple inequality will often be found useful.

4. Another important biometric function is the *expectation of life* $\bar{e}_x$, defined by

$$\bar{e}_x = \int_0^\infty {}_t p_x dt = \frac{1}{l_x} \int_x^\infty l_x dx. \qquad \ldots\ldots(10)$$

We may also, by (3), express the expectation of life in terms of μ_x, thus

$$\bar{e}_x = \int_0^\infty e^{-\int_x^{x+t} \mu_x dx} dt. \qquad \ldots\ldots(11)$$

Inserting the lower limit to μ_x by (1), we obtain, on performing the integration,

$$\bar{e}_x \leqslant \frac{1}{\epsilon}. \qquad \ldots\ldots(12)$$

That is, there exists a boundary, independent of x, which $\bar{e}_x$ cannot exceed. This result is trivial, but a more important inequality may be proved, if it is assumed that μ_x *never decreases for* $x \geqslant x_0$. In that case, it follows from (11) that

$$\bar{e}_x \leqslant \int_0^\infty e^{-t\mu_x}\, dt,$$

or

$$\bar{e}_x \leqslant \frac{1}{\mu_x} \qquad (x \geqslant x_0), \qquad \text{......(13)}$$

whence, by (9),

$$\bar{e}_x < \frac{1}{q_{x-1}} \qquad (x \geqslant x_0 + 1). \qquad \text{......(14)}$$

Instead of $\bar{e}_x$ it is often sufficient to consider the *curtate expectation of life* e_x defined by

$$e_x = \sum_{t=1}^{\infty} {}_tp_x = \frac{1}{l_x} \sum_{t=1}^{\infty} l_{x+t}, \qquad \text{......(15)}$$

which, as l_x is constantly decreasing, is always smaller than $\bar{e}_x$. We evidently have

$$e_x = p_x + p_x p_{x+1} + p_x p_{x+1} p_{x+2} + \dots,$$

so that, if p_x *does not increase with* x,

$$e_x \leqslant p_x + p_x^2 + p_x^3 + \dots,$$

or

$$e_x \leqslant \frac{p_x}{q_x}. \qquad \text{......(16)}$$

From (15) we immediately find

$$e_x = p_x(1 + e_{x+1}). \qquad \text{......(17)}$$

This relation shows that if $q_x \to 1$ for $x \to \infty$, then $e_x \to 0$ for $x \to \infty$, and conversely: if $e_x \to 0$, then $q_x \to 1$.

It follows that if $\bar{e}_x \to 0$, then $\mu_x \to \infty$. The converse proposition may be proved thus: If $\mu_x \to \infty$ as $x \to \infty$, we may choose x so large that $\mu_x > K$ for all values of x above a certain limit, and consequently by (11), $\bar{e}_x < \frac{1}{K}$ for all values of x above that limit. But K being arbitrary, we have $\bar{e}_x \to 0$ if $\mu_x \to \infty$.

5. What the text-books say about the function l_x is as a rule not expressed with sufficient reserve, and is even in many cases positively misleading. Without going into details about the various forms of vagueness or inaccuracy I have observed, I think it safe to assert that an ordinary student reading a text-book for the first time may be led to form the following opinions on the nature of the table of l_x.

Assume that l_0 persons are born, and that we follow these persons from their birth till their death, the number of those who are alive after x years being denoted by l_x. They will certainly all die within a finite time. There is therefore an "oldest age" ω such that $l_x > 0$ for $x < \omega$, and $l_x = 0$ for $x \geqslant \omega$. No question of convergence arises, for we have simply

$$\int_x^\infty l_x dx = \int_x^\omega l_x dx,$$

the limits of integration being in reality finite. If an analytical expression is assumed for l_x, it may safely be assumed that $\int_\omega^\infty$ is so small that it can be neglected.

Our student is confirmed in this view when he discovers that the table of l_x is stated in integral numbers, commencing with, say, 100,000 persons at the lowest age and terminating with 0 at an age about 100 years.

But this point of view contains a fallacy. It is true that any given *finite* number of persons must all die within a finite time or, to put it more precisely, the probability that one or more of them will survive indefinitely is zero. This follows from the above established fact that ${}_tp_x \to 0$ when $t \to \infty$. But the greater that number of persons is, the greater will also be the number of them who may still be alive at age 100 or any other assigned age, and it is therefore quite impossible to maintain the existence of any definite "oldest age" ω within which everybody must certainly die. Let us, to put it mathematically, assume for a moment that ω is the upper limit of the age which human life can attain. Then we have admitted that it is possible that a person may be alive at the age $\omega - \eta$, where η is a quantity which we may choose *as small as we please*, for instance equal to one second. But at the exact age ω, or one second later, that person must, according to hypothesis, necessarily be dead. Does anybody really believe that there is an age ω with this miraculous peculiarity?

It might, however, be argued that in introducing the function l_x in the way we have advocated above, we only replace one monstrosity by another, for we admit the possibility of a person being alive at any age, however advanced. Nobody will ever believe that a person can live to become 1000 years old, and from a practical point of view this may safely be maintained. It is *practically* immaterial whether we say that a person *cannot* attain the age of 1000 years, or that the *probability* of attaining that age* is $< 10^{-10^{35}}$, an inconceivably small number. But if

* According to the $O^{M(5)}$ table as graduated by Makeham's formula we have

$$\log l_x = 5{\cdot}0575047 - {\cdot}0025575x - 10^{{\cdot}039x + \bar{4}{\cdot}7007037},$$

whence, for instance,

$$\log \frac{l_{1000}}{l_{10}} < -10^{35}.$$

theoretical clearness can be gained by speaking of extremely small probabilities instead of impossibilities, it should certainly be done. Infinity is not a number, but only denotes the absence of a boundary, and the non-occurrence in practice of observations of a certain order of magnitude should not without necessity be attributed to a mysterious boundary hidden somewhere, as it is always sufficiently accounted for by their exceedingly small probabilities, exactly as in the Theory of Errors. Even in such an everyday subject as the Theory of Interest we do not hesitate to make use of infinite durations in cases where no boundary can be indicated, not because anybody believes that a perpetuity will really continue to be paid *ad infinitum*, but because it is a convenient and harmless construction, as the very distant payments do not appreciably influence the value of the perpetuity.

6. An entirely different point of view from the one we have discussed is that in actually constructing a table of l_x it is necessary to stop at a certain age which we may still call ω, although it has nothing to do with an "oldest age" in the abstract sense of the word. Our ω will now be determined, not by an imagined impossibility of living beyond that age, but by considering the error committed in calculating integrals and sums to the limit ω instead of to infinity.

It seems reasonable to determine ω in such a way that in calculating the expectation of life by the approximate formula

$$\bar{e}_x = \frac{1}{l_x}\int_x^\omega l_x dx \qquad \text{......(18)}$$

we obtain at least three reliable decimals in the result. In that case we must have

$$\frac{1}{l_x}\int_\omega^\infty l_x dx \leqslant \cdot 0005.$$

Assuming now that μ_x does not decrease for $x > \omega$, we have according to (13)

$$\int_\omega^\infty l_x dx < \frac{l_\omega}{\mu_\omega},$$

so that we shall have three correct decimals in $\bar{e}_x$ provided that

$$\frac{1}{l_x}\cdot\frac{l_\omega}{\mu_\omega} \leqslant \cdot 0005,$$

that is, if
$$l_x \geqslant 2000\,\frac{l_\omega}{\mu_\omega}.$$

It follows from this that we shall have at least three correct decimals in $\bar{e}_x$ calculated by (18), when $x \leqslant 100$, if ω is calculated by

$$l_{100} = 2000 \frac{l_\omega}{\mu_\omega}. \qquad \text{......(19)}$$

In the $O^{M(5)}$ table as graduated by Makeham's formula we have

$$\mu_x = \cdot 0058889 + 10^{\cdot 039x + \bar{4}\cdot 0161709};$$

by this and by the expression for $\log l_x$ given above we find $106 < \omega < 107$. The table of l_x must therefore be prolonged as far as 107 if we want the expectation of life with three correct decimals for all ages up to 100 years, neglecting the contribution from ages above 107. And—this is material—l_x should be given with a sufficient number of figures throughout, say five or six significant* figures, and not, as at present, finishing off with $l_{100} = 7$, $l_{101} = 3$, $l_{102} = 1$.

There are, of course, other ways of determining ω; we may, for instance, consider sums instead of integrals; or we may fix ω arbitrarily, in which case it would be necessary to investigate how many decimals in the expectation of life may be relied upon at the various ages. But in all cases the principle remains the same, and we need hardly go into further details.

It is sometimes objected that there is no meaning in stating l_x and similar functions with five significant figures at the highest ages, because the observations available there, even after graduation, are insufficient for producing this degree of accuracy, and also because the table is to be applied to the future which never agrees wholly with the past.

This question is closely connected with the question of the purpose of graduation. It may be said in a general way that the object of graduation is to replace the rough observations by a more "smooth" series of data; but the smoothness thus obtained is partly lost again, if the last few values of l_x are only stated with one or two significant figures. If now we proceed one step further and ask *why* we want the table to be smooth, the answer is: Not only, because a smooth table is likely to be closer to the truth than an irregular one, but chiefly because we want to be justified in applying methods of interpolation, numerical differentiation and integration, etc., in short *mathematical methods*, to the table. But this requires that the function represented by the table possesses a differential coefficient of a certain order, and the simplest way to ensure this is to graduate

* The "significant" figures commence with the first figure that is different from zero.

the table by an analytical function. Under these circumstances there is evidently a need for retaining a not too small number of figures throughout the table, and this number of figures has nothing to do with the accuracy of the observations. It depends on the use that is to be made of the table, and remains the same whether the table is a purely hypothetical one, or has been derived from a very large number of observations.

7. As an application of the principles discussed above, we will consider the classical problem: whether it is possible to have two different mortality tables producing, for all ages at entry and durations, the same policy values for a life assurance with annual premiums.

Denoting, as usual, the life annuity-due by a_x, and the policy value after t years by ${}_tV_x$, we have*

$$ {}_tV_x = 1 - \frac{\mathrm{a}_{x+t}}{\mathrm{a}_x}. \qquad \text{......(20)} $$

Let μ_x^{I} be the force of mortality according to another table of mortality; the corresponding life-annuities, etc., will be denoted by $\mathrm{a}_x^{\mathrm{I}}$, etc.

If, now, we are to have ${}_tV_x = {}_tV_x^{\mathrm{I}}$, we must, according to (20), have

$$ \frac{\mathrm{a}_{x+t}}{\mathrm{a}_{x+t}^{\mathrm{I}}} = \frac{\mathrm{a}_x}{\mathrm{a}_x^{\mathrm{I}}} \qquad \text{......(21)} $$

for all values of x and t. Denoting by k a constant, the condition (21) may therefore be written

$$ \frac{\mathrm{a}_x}{\mathrm{a}_x^{\mathrm{I}}} = 1 - k. \qquad \text{......(22)} $$

It is obvious that $k < 1$, as the expression on the left is always positive. But k must as a rule also satisfy another condition. Let us assume that the original mortality table has been graduated by such a formula (for instance Makeham's formula) that $\mu_x \to \infty$ as $x \to \infty$. This is, according to paragraph 3, equivalent to saying that $q_x \to 1$ as $x \to \infty$. We therefore have $\mathrm{a}_x \to 1$ as $x \to \infty$, as follows from the obvious relation

$$ \mathrm{a}_x = 1 + v p_x \mathrm{a}_{x+1}, \qquad \text{......(23)} $$

where v is the present value of a unit, due one year hence. But if $\mathrm{a}_x \to 1$, it follows from (22) that $\mathrm{a}_x^{\mathrm{I}}$ tends to a limiting value that *exceeds* 1 (as $k < 1$). For $k = 0$ would mean that the two

* *Institute of Actuaries' Text-Book*, Part II, Second Ed. (1902), p. 323.

mortality tables, against hypothesis, were identical, and $k < 0$, that a_x^I could become smaller than unity. We therefore have $0 < k < 1$.

Let us now put

$$\lim_{x \to \infty} a_x^I = 1 + \eta \qquad (\eta > 0). \qquad \text{......(24)}$$

In the equation

$$a_x^I = 1 + v p_x^I a_{x+1}^I \qquad \text{......(25)}$$

we let $x \to \infty$ and introduce (24), putting $v = \frac{1}{1+i}$, where i is the rate of interest. We then have

$$1 + \eta = 1 + \frac{1+\eta}{1+i}(1 - q_\infty^I),$$

whence

$$q_\infty^I = \frac{1 - \eta i}{1 + \eta}. \qquad \text{......(26)}$$

A mortality table producing such annuity values a_x^I that (22) is satisfied must therefore be a peculiar one, as $q_\infty^I < 1$. It is easy to construct the table, for we obtain from (25)

$$p_x^I = \frac{a_x^I - 1}{v a_{x+1}^I},$$

whence, as

$$a_x^I = \frac{a_x}{1-k} \quad \text{and} \quad a_{x+1}^I = \frac{a_{x+1}}{1-k},$$

$$p_x^I = \frac{a_x - 1 + k}{v a_{x+1}}, \qquad \text{......(27)}$$

or, eliminating a_{x+1} by (23)

$$p_x^I = p_x\left(1 + \frac{k}{a_x - 1}\right). \qquad \text{......(28)}$$

This suffices for constructing the table of l_x^I. But we still have to satisfy ourselves that it is possible to give k such a value, comprised between 0 and 1, that all the values of p_x^I obtained from (28) are confined to the interval from 0 to 1, as otherwise they could not represent probabilities. Now, as k is positive, p_x^I is evidently always positive. The condition $p_x^I < 1$ can be written

$$k < \frac{q_x}{p_x}(a_x - 1),$$

or by (23)

$$k < v q_x a_{x+1}. \qquad \text{......(29)}$$

This inequality must be satisfied for all values of x, and besides we must have $0 < k < 1$. It is easy to see that values of k exist