TRANSACTIONS OF SOCIETY OF ACTUARIES
1969 VOL. 21 PT. 1 NO. 59 AB

A LOGICAL APPROACH TO POPULATION PROBLEMS

ROBERT W. BATTEN

ABSTRACT

The student of life contingencies has been struggling with the various types of stationary population problems for many years. Many approaches to such problems have been suggested, ranging from the classical integration method to clever, well-founded diagrammatic methods. Perhaps much of the prevailing confusion among students is attributable to the failure to master any single technique.

The purpose of this paper is to present an approach to such problems that is general in nature in that it requires neither diagrams nor integration—only a clear understanding of the basic symbols and concepts involved. It is felt that any actuarial problem, be it practical or purely theoretical, should be handled by powers of general reasoning and logical thought processes wherever feasible. Problems involving stationary populations are among the few types of actuarial problems which can be generally handled by appealing only to basic principles.

By considering first the simpler, then the more complex problem types, this paper applies the basic demographic concepts to problems similar to those posed by the Part 4 Examination Committee in recent years. A final portion of the paper deals with a limited application of these principles to situations involving continuously increasing birth rates.

I. INTRODUCTION

The actuarial student has historically been the beneficiary of clever devices designed by his predecessors to simplify particular problem types with which he must wrestle. The late Harry Gershenson was probably the student's closest ally, since his "toll-road" concept of mortality measurement greatly simplified a previously nightmarish topic for students. He also made significant contributions to students' understanding of many concepts in life contingencies and graduation.

In some instances, however, it is the feeling that certain short-cut devices have worked against the actuarial student's better interests. If a student finds himself able to solve certain problems with an entirely mechanical process, the natural result often is unwillingness to investigate the theoretical aspects of the situation, thus defeating the purposes of the Society examinations.

Problems involving stationary populations have long been considered among the most difficult to prospective actuaries. The classical method, involving complicated integrals, was long ago recognized as both tedious and highly subject to careless error. An article in the *1950 Transactions* authored by Mr. Walter L. Grace and Dr. Cecil J. Nesbitt represented a first attempt to provide a simpler general method of solution. In the *1964 Transactions*, Mr. Kenneth Veit outlined a diagrammatic approach in which he enlarged upon the previously espoused Grace-Nesbitt principles. Mr. Veit's method was quite ingenious and has doubtless been a saving grace for many life contingencies students during the last four years.

The previous contributions to the actuarial literature mentioned above were made by highly competent actuaries and can be backed up by rigorous mathematical theory, but the author's view is that this subject area can be handled with much less ado. The contention forwarded here is that *any* problem involving a stationary population concept can be solved, without integrals, diagrams, or mnemonic devices, in a comparatively short time. This approach involves only general reasoning and full understanding of the basic characteristics of a stationary population. A person who is not mathematically inclined should find this method quite comprehensible, and the student plagued by a lack of mathematical agility should find himself in a much more competitive position than before.

It should be made clear that, although this paper was written primarily for the actuarial student of life contingencies, it could also be helpful to demographers in population projections.

II. REVIEW OF BASIC DEFINITIONS

Since our nonmathematical approach will involve only logical applications of basic definitions, it will be helpful to state these definitions.

l_x = Number of persons that *attain* exact age x during *any* one-year period.

T_x = Number of persons *attaining* age x or over at *any instant* of time

or

Aggregate future lifetime (from age x until death) of those l_x persons who attain age x in a one-year period.

Y_x = Aggregate *future* lifetime of those T_x persons now age x and over

or

Aggregate past lifetime, *since age x*, of those T_x persons now age x or over.[1]

[1] C. Wallace Jordan, *Life Contingencies* (Chicago: Society of Actuaries, 1952, pp. 247–48).

III. A BASIC THEOREM

There is a particular result that has been stated in actuarial literature but for which a proof does not seem to have been precisely formulated. This theorem is of major importance in the development of this paper.

Let us consider the $T_x - T_{x+n}$ persons now between ages x and $x + n$ in a stationary population. From our basic knowledge of stationary population concepts, we know that $n \cdot l_{x+m}$ of these persons will ultimately attain age $x + m$, if $m \geq n$. Further, we will prove the following statement:

The average present attained age of those persons now living between ages x *and* x + n *who will ultimately attain age* x + m, m ≥ n, *is exactly* x + ½n.

A mathematical approach is useful for verifying this assertion. The total past lifetime of those persons now between ages x and $x + n$ may be expressed by

$$\int_{t=0}^{t=n} (x + t) l_{x+t} dt ,$$

which may easily be shown to be

$$x \cdot T_x - (x + n) T_{x+n} + Y_x - Y_{x+n} .$$

From this quantity we will subtract the past lifetime of those persons with whom we are not concerned, that is, the $T_x - T_{x+n} - n \cdot l_{x+m}$ persons who do not survive to age $x + m$.

This quantity can be expressed by

$$\int_{y=x}^{y=x+n} \int_{t=0}^{t=x+m-y} y \cdot l_{y+t} \, \mu_{y+t} \, dt dy$$

which may be shown to be

$$x \cdot T_x - (x + n) T_{x+n} + Y_x - Y_{x+n} - l_{x+m}(n \cdot x + \tfrac{1}{2} n^2) .$$

The difference between these two quantities is clearly the past lifetime of the $n \cdot l_{x+m}$ persons in whom we are interested and is

$$l_{x+m}(n \cdot x + \tfrac{1}{2} n^2) .$$

If we divide by the number of persons involved, the average past lifetime of this group is shown to be exactly $x + \frac{1}{2}n$.

As a result of this theorem, we can now determine, for example, that the average attained age of those persons now between ages 25 and 40 who will survive to age 65 (or to any other age greater than or equal to 40) is precisely thirty-two and one-half years.

We are also able to comment upon the average age of those $T_x -$

$T_{x+n} - n \cdot l_{x+m}$ persons in this group who will die *before* attaining age $x + m$. Since the average age of persons now alive between any two given ages x and $x + n$ is clearly less than $x + \frac{1}{2}n$, an application of the result proved above allows us to assert that the $T_x - T_{x+n} - n \cdot l_{x+m}$ persons must be younger, on the average, than $x + \frac{1}{2}n$.

The above results and a full understanding of the meaning of the various stationary population symbols are now our only available tools. With them we will be able to find expressions for any desired quantities, such as past or future lifetime of specific groups of individuals, as well as many variations thereof. We will now investigate properties of the various groups of lives which are of particular interest in analyses of stationary populations.

IV. ILLUSTRATIONS OF THE GENERAL REASONING APPROACH

1. *Persons Who Die between Two Given Ages*

The first body of lives that we will consider will be those persons who will die between two given ages, say, x and $x + n$, in a given year. These people are $l_x - l_{x+n}$ in number, l_x being the number of lives attaining exact age x during the year and l_{x+n} the number attaining exact age $x + n$ during the same year. In order for the population between these two ages to remain stationary, the excess of l_x over l_{x+n} must exit by death.

Let us consider the total lifetime of the $m(l_x - l_{x+n})$ persons who will die in the next m years between the ages of x and $x + n$. Clearly each of these persons must live at least x years, but, in addition, each will have some lifetime *after* attainment of age x. It will be convenient to consider such lifetime of the $m \cdot l_x$ lives and the $m \cdot l_{x+n}$ lives separately.

The $m \cdot l_x$ lives have an aggregate lifetime, after age x, of $m \cdot T_x$ years. The lifetime, after age x, of the $m \cdot l_{x+n}$ lives must be considered in two steps. First, each person in this group must, by definition, survive the entire n years between x and $x + n$. Further, the group will live a total of $m \cdot T_{x+n}$ years after attainment of $x + n$.

Combining terms, we see that the total lifetime of these $m(l_x - l_{x+n})$ lives is

$$x \cdot m(l_x - l_{x+n}) + m \cdot T_x - (m \cdot T_{x+n} + m \cdot n \cdot l_{x+n}) .$$

Dividing this by the number of persons involved gives us the familiar result for their *average* lifetime:

$$x + \frac{T_x - T_{x+n} - n \cdot l_{x+n}}{l_x - l_{x+n}}$$

Notice that the fact that the problem covered an m-year period had no effect on the average lifetime, but the aggregate lifetime reflects the fact that m one-year groups are covered by the problem.

2. Persons Now Alive between Two Given Ages

Another group of lives often encountered consists of those persons between two given ages, say, x and $x + n$, at the inception of a given problem. We know that there are $T_x - T_{x+n}$ such persons. Let us consider their total lifetime.

As before, $x(T_x - T_{x+n})$ represents the years lived by this group before attainment of age x. The rest of their lifetime, as in the preceding example, can best be determined by first considering the T_x persons and then adjusting our result by the lifetime of the T_{x+n} persons who are not included in the statement of the problem.

As for the T_x persons, our definitions tell us that this group has a past lifetime, since age x, of Y_x years and a future lifetime of an additional Y_x years.

Similarly, the T_{x+n} persons have a past lifetime since age $x + n$ of Y_{x+n} years and a future lifetime of an additional Y_{x+n} years. But we must not forget that each member of this group lived the entire n-year period between x and $x + n$, creating an additional $n \cdot T_{x+n}$ years. Hence the total past lifetime of the T_{x+n} persons, after attainment of age x, is $Y_{x+n} + n \cdot T_{x+n}$ years.

The total lifetime of the $T_x - T_{x+n}$ lives can then be allocated as follows:

Lifetime before age x:	$x(T_x - T_{x+n})$;
Past lifetime since x:	$Y_x - Y_{x+n} - n \cdot T_{x+n}$;
Future lifetime:	$Y_x - Y_{x+n}$.

Proper combination of these quantities, followed by division by $T_x - T_{x+n}$, will result in the group's average age, average age since x, or average age at death. The average expectation of life for a person in this group is easily seen to be

$$\frac{Y_x - Y_{x+n}}{T_x - T_{x+n}} .$$

Finally, suppose we desire the average future lifetime of this group before attainment of age $x + m$, $m \geq n$. First, the T_x persons have an aggregate future lifetime of Y_x years. Similarly, the T_{x+n} persons have a future lifetime of Y_{x+n} years. From these $Y_x - Y_{x+n}$ years, however, we must subtract those years to be lived *after* attainment of $x + m$. This is not difficult, as we know that $n \cdot l_{x+m}$ persons will survive to age $x + m$, and the lifetime, after $x + m$, of these persons is $n \cdot T_{x+m}$. Our final result is, then,

$$\frac{Y_x - Y_{x+n} - n \cdot T_{x+m}}{T_x - T_{x+n}}.$$

3. *Persons Alive between Two Given Ages Who Will Die within* m *Years*

Suppose we are concerned with those persons now between ages x and $x + n$ who will die within the next m years. Clearly the number who will *not* die is $T_{x+m} - T_{x+m+n}$; thus the persons whom we must consider are $(T_x - T_{x+n}) - (T_{x+m} - T_{x+m+n})$ in number.

Let us consider the average present age of this group. The past lifetime, since age x, of the $T_x - T_{x+n}$ persons is $Y_x - Y_{x+n} - n \cdot T_{x+n}$ years, as previously discussed. From this quantity, we need to subtract the past lifetime, since age x, of the $T_{x+m} - T_{x+m+n}$ survivors, not as of the "end" of the problem but as of the "beginning." In order to do this, we must calculate their past lifetime, since age x, as of the "end" of the problem and reduce this by the m years which each survivor lived during the course of the m-year period. This results in a past lifetime, since age x, at problem inception, of

$$Y_{x+m} - Y_{x+m+n} - n \cdot T_{x+m+n} + m(T_{x+m} - T_{x+m+n}) - m(T_{x+m} - T_{x+m+n})$$

or merely

$$Y_{x+m} - Y_{x+m+n} - n \cdot T_{x+m+n} \text{ years.}$$

Collecting terms, we find that the average age of this over-all group is

$$x + \frac{(Y_x - Y_{x+n}) - (Y_{x+m} - Y_{x+m+n}) - n(T_{x+n} - T_{x+m+n})}{(T_x - T_{x+n}) - (T_{x+m} - T_{x+m+n})}.$$

4. *Persons Alive between Two Given Ages Who Will Not Survive to a Given Higher Age*

Finally, let us consider those persons now between ages x and $x + n$ who will not survive to age $x + m$, $m \geq n$. It has been previously determined that these persons are $T_x - T_{x+n} - n \cdot l_{x+m}$ in number, and it should be clear that the investigation of this group of lives will be quite similar to that followed in Section 2.

Let us begin by determining the future lifetime of this group of persons. The $T_x - T_{x+n}$ persons clearly have an aggregate future lifetime of $Y_x - Y_{x+n}$ years, but this must be reduced by the future lifetime of the $n \cdot l_{x+m}$ persons excluded from consideration in this problem. Applying our basic theorem, we recall that these persons may be assumed to be exact age

$x + \frac{1}{2}n$. By the very definition of this group, we are certain that each member will survive to $x + m$. Hence their future lifetime will be $(m - \frac{1}{2}n)n \cdot l_{x+m}$ years in addition to the $n \cdot T_{x+m}$ which they, as a group, will live after attaining age $x + m$. Collecting terms, we see that the desired quantity is

$$Y_x - Y_{x+n} - [(m - \tfrac{1}{2}n)n \cdot l_{x+m} + n \cdot T_{x+m}] .$$

Similarly, the past lifetime of this group, *since age x*, is

$$Y_x - Y_{x+n} - n \cdot T_{x+n} - \tfrac{1}{2} \cdot n^2 \cdot l_{x+m} .$$

Hence, the total lifetime of these $T_x - T_{x+n} - n \cdot l_{x+m}$ persons may be categorized as follows:

Lifetime before age x: $x(T_x - T_{x+n} - n \cdot l_{x+m})$;

Past lifetime since x: $Y_x - Y_{x+n} - n \cdot T_{x+n} - \frac{1}{2}n^2 \cdot l_{x+m}$;

Future lifetime: $Y_x - Y_{x+n} - [(m - \frac{1}{2}n)n \cdot l_{x+m} + n \cdot T_{x+m}]$;

Total: $x(T_x - T_{x+n} - n \cdot l_{x+m}) + 2(Y_x - Y_{x+n})$

$$- n(T_{x+n} + T_{x+m}) - m \cdot n \cdot l_{x+m} .$$

This result could be verified by solution of the following integral:

$$\int_{y=x}^{y=x+n} \int_{t=0}^{t=x+m-y} (y + t) l_{y+t} \mu_{y+t} \, dt dy .$$

It is important to realize that, since we solved this problem logically rather than mechanically and since the answer is broken into its component parts, we are able to answer any of a number of questions concerning this group. For example, the average age at death, average attained age, and average life expectancy are now immediately obtainable.

V. A FINAL ILLUSTRATION

In order to show the value of a logical approach, let us consider a final problem that illustrates many of the points discussed. A student with a full understanding of this paper should be able to write down the solution in a very short time.

Find an expression for the average attained age of those persons in a stationary population now between 25 and 40 who will die between the ages of 30 and 50 within the next 20 years.

Our first step must be, as always, to find the number of persons involved. A general rule to follow is to include those persons who become "eligible" to die within the restraints of the problem and then to subtract those persons when, due to attainment of a certain age or passage of a certain period of time, their deaths would no longer be of pertinence to the given problem.

In this case, those persons originally between 25 and 30 do not become "eligible" to die until they attain age 30, while those originally between 30 and 40 are eligible at the outset of the problem. It should be clear that the $T_{25} - T_{30}$ persons must be considered for 20 years, while the $T_{30} - T_{40}$ persons are followed until they "age out" of the problem at age 50. Therefore, of the $T_{25} - T_{30}$ persons, we must consider

$$5 \cdot l_{30} - (T_{45} - T_{50}) , \quad (1)$$

and of the $T_{30} - T_{40}$ persons,

$$T_{30} - T_{40} - 10 \cdot l_{50} . \quad (2)$$

Notice that the 20-year restriction did not apply to the $T_{30} - T_{40}$ persons and that the "age 50" restriction had no effect on the $T_{25} - T_{30}$ persons. Once it is seen that age 30 is the convenient breaking point for the $T_{25} - T_{40}$ lives, it becomes very easy to determine the number of lives to include in the denominator. The student should realize that it is often necessary to consider more than just two subgroups of persons in order to obtain the number of deaths to be included.

For Group 1 the aggregate past lifetime may be subdivided, using the methods of this paper, as follows:

Lifetime before age 25: $25[5 \cdot l_{30} - (T_{45} - T_{50})]$; (3)

Lifetime since age 25: $2\frac{1}{2}(5 \cdot l_{30}) - (Y_{45} - Y_{50} - 5 \cdot T_{50})$. (4)

Group 2 may be handled similarly:

Lifetime before age 30: $30[T_{30} - T_{40} - 10 \cdot l_{50}]$; (5)

Lifetime since age 30: $(Y_{30} - Y_{40} - 10 \cdot T_{40}) - 5(10 \cdot l_{50})$. (6)

In quantity (6), it should be clear that the $10 \cdot l_{50}$ persons who survive to age 50 are assumed to be exact age 35 at the inception of the problem, since they are survivors of an original group between ages 30 and 40. A similar assumption was used to derive quantity (4).

Finally, then, the answer to our problem must be

$$\frac{(3) + (4) + (5) + (6)}{(1) + (2)} .$$

VI. INCREASING POPULATIONS

The preceding sections of this paper have been concerned only with the classical stationary population, or a population in which the birth and death rates remain unchanged from year to year. In reality, of course, most populations are not stationary, and it would therefore be quite advantageous to be able to apply the above theory to a fluctuating population with, say, a continuously increasing birth rate. Unfortunately, the mathematics of such an investigation does not seem to lend itself to the simple, concise symbols and formulas that would be desirable. Nonetheless, it might be of interest to discuss a few pertinent results.

Let us assume that the number of births in a community increases continuously at the constant annual rate δ. Let us further assume that the community is not subject to migration and that the mortality rates at each age remain constant, as they do in a population that is stationary.[2] A first problem might be to determine the size of the total population aged x and over at some given time, corresponding to the T_x of the stationary population. It will be assumed that the birth increase has been in effect for a number of years greater than the terminal age of the mortality table.

Of first priority is the determination of the annualized number of births in effect exactly t years after the increase has been operative. Since the number of births is increasing continuously at a constant annual rate δ, this annualized number of births at time t must be $e^{\delta t}R_0$, where R_0 is the number of births per year before the increase was initiated. The number of persons attaining some random age y is clearly dependent upon the number of births exactly y years before. Thus, at time t, the number of persons living between ages x and $x + 1$ could be expressed as

$$\int_{y=x}^{y=x+1} e^{\delta(t-y)} l_y dy \, ,$$

which may be easily shown to be

$$(1 + i)^t \int_{y=x}^{y=x+1} v^y l_y dy \, ,$$

where $i = e^\delta - 1$ and $v = e^{-\delta}$. It is now clear that the total population aged x and over at time t under the new condition is $(1 + i)^t \bar{N}_x$.

We are able to apply our general reasoning approach to the above integral expression in order to find the aggregate future lifetime of those $(1 + i)^t \bar{N}_x$ persons age x and over at time t. The $v^y l_y dy$ persons attaining age y clearly have an aggregate future lifetime of $v^y T_y dy$ years; hence the aggregate figure desired is

$$\int_{y=x}^{y=\infty} v^y T_y dy \, .$$

[2] It does not seem feasible to apply the logical approach outlined in this paper to populations in which mortality rates do not remain the same from year to year.

The total past lifetime since age x of these $(1+i)^t\bar{N}_x$ lives is determined by crediting each of the $l_y dy$ lives in the integral expression with $(y-x)$ years, resulting in

$$(1+i)^t \int_{y=x}^{y=\infty} v^y (y-x) l_y dy$$

or

$$(1+i)^t D_x (\bar{I}\bar{a})_x .$$

Combining all of the above results, we see that the average age at death of those persons aged x and over is

$$x + \frac{\int_{y=x}^{y=\infty} v^y T_y dy}{\bar{N}_x} + \frac{(\bar{I}\bar{a})_x}{\bar{a}_x} ,$$

a result which is shown on the 1956 Part 4 Examination of the Society of Actuaries. Notice that this average age is independent of the time at which the population is examined.

Investigations of other groups of lives may now be made, using the principles shown above and the concepts outlined in the first few sections of this paper. Unfortunately, however, the average age of those persons now between ages x and $x+n$ who survive to $x+m$, $m \geq n$, is not equal to $x+\frac{1}{2}n$ when we are dealing with a variable group. In any investigation which is made, one should always be careful to verify that his answers, evaluated at $i=0$, are equal to those results previously found for the standard stationary population.

VII. SUMMARY

This paper has stated and verbally derived most of the important results that the actuarial student needs in his quest for a mastery of population problems. However, it is highly recommended that each individual problem be attacked from basic principles. Only in this way may a student gain confidence in himself and develop a full understanding of the significance of each symbol. In the long run, he will find that the time necessary to handle complicated problems will decrease as his familiarity with the important logical thought processes increases.