STATIONARY POPULATION PROBLEMS

David C. Jesionowski

WHY ANOTHER MODEL?

Certainly one of the most difficult subjects for an actuarial student to comprehend is the stationary population. To prepare for examination questions concerning this perplexing topic, I studied chapter eight of Jordan and I read various T.S.A. articles and the subsequent discussions which followed each article. I benefited from reading all of the material; however, the material written by Mr. Veit, Dr. Nesbitt, and Mrs. Tino proved to be the most beneficial.

From the T.S.A. article written by Mr. Veit, I learned how to apply the "in-and-out" method to solve a variety of stationary population problems. Dr. Nesbitt's discussion of Mr. Veit's article revealed that "the total number of persons aged x or more at any time, is made up of infinitesimal survivorship groups, a typical one consisting of $l_y dy$ persons between ages y and y+dy"[1] and this led me to look one step further. Finally, Mrs. Tino's discussion of Mr. Veit's article concerned the geometrical interpretation of the stationary population and provided insight into the reasons why the "in-and-out" method worked. Despite the fact that I had studied all of this material, I did not have a complete understanding of the stationary population. Because the time before the exam was getting short,

[1] T.S.A., 1965, XVI, pg. 244

I decided not to study the stationary population any further and I concentrated my efforts on the remaining topics.

In the course of my review of the stationary population material prior to the exam, I noticed on aspect of the present model (views depicted by figures 1a and 1b)[2] which I could not justify.

Figure 1a

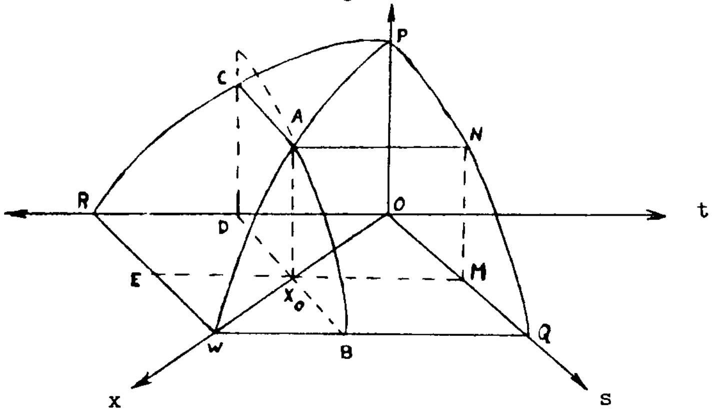
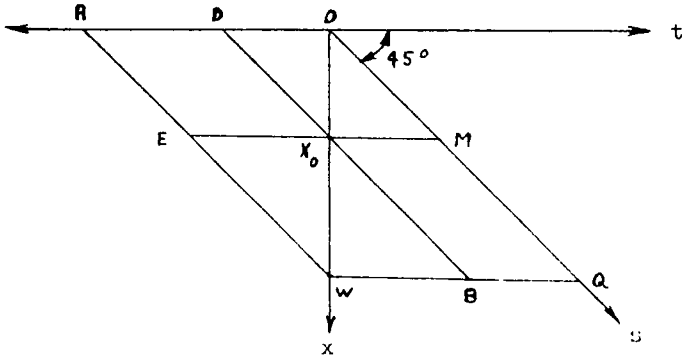


Figure 1b



[2]Models 1a and 1b are shown as figures 4 and 5, respectively, in Paulette Tino's discussion of Mr. Veit's article. Vector s, not shown as such in Mrs. Tino's material, has been added by the author. T.S.A., 1965, XVI, pg. 256

(3)

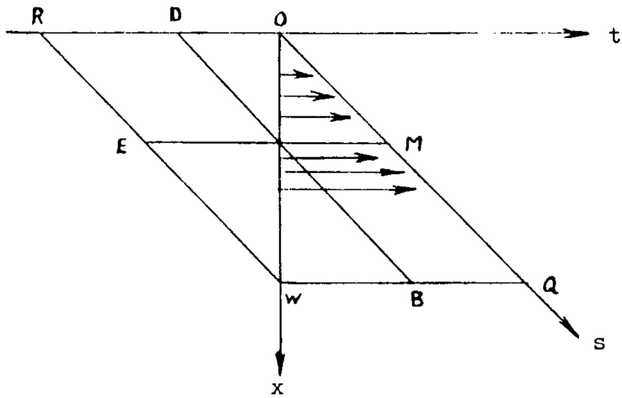
The fact that I could not justify one aspect of the present model certainly does not indicate that there is an underlying mathematical fallacy. More likely, it is indicative of my mathematical maturity.

Before discussing the problem which I encountered, it is necessary to define specific parts of the model and to note several properties. For purposes of our discussion, I would like to define axis t as the "time of study" axis, and axis x as the "ages of entry into the study" axis. The properties of the model which should be noted are:

(1) Δ OwQ is an isoceles right triangle
(2) $\|Ow\| = \|wQ\| < \|OQ\|$
(3) $\|Ow\|\cdot\sqrt{2} = \|OQ\| = w\sqrt{2}$
(4) Above $\overline{Ow}$, one would find an l_x curve with the height at point O being l_0
(5) If one integrates with respect to x from point O to point w, the area under the l_x curve would be found to be T_0

With these definitions and properties in mind, let us examine the problem which I encountered when I attempted to use the present model to determine the total future lifetime of a group of lives now age 0. The total future lifetime of such a group is T_0 and I felt that the area under the curve above vector s should therefore be T_0. As shown in figure 2, if I integrated with respect to t (time of study) over all values of x (ages of entry into the study), I would cross $\overline{OQ}$ in such a manner as to obtain the correct area, T_0. However, I could not justify the need to cover all values of x, in order to obtain the proper area.

Figure 2



Next, I attempted to integrate (as I understand integration) over vector s. Initially I believed the area would be T_0 . I later noted that $\|Ow\|$ and $\|OQ\|$ were not equivalent and I was not able to understand how one could obtain identical areas beneath the curves above two line segments of different length.

The actual situation seemed far too complex; so, I simplified it by defining s(x) as follows:

$$s(x) = \begin{cases} 1, & \text{for } 0 \leq x < w \\ 0, & \text{for } x \geq w \end{cases}$$

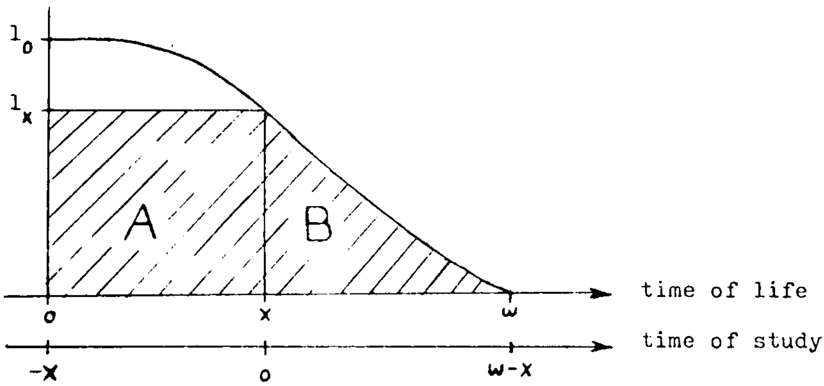
In such a situation, every member of the population survives to age w, the limiting age. Clearly, $l_x = l_0$ for $0 \leq x < w$. The area under this l_x curve from point 0 to point w, would be $w \cdot l_0$. However, if I integrated with respect to s from point O to point Q, I found the area to be $w \cdot l_0 \cdot \sqrt{2}$.

Because of this enigma, I searched for a new model. I wanted a model in which line segments $\overline{Ow}$ and $\overline{OQ}$ were of equal length and which could be used to solve any stationary population problem (including finding the total future lifetime of those now living between ages x and x+t, who die before age x+t). After several days of research, I found such a model.

DEVELOPMENT AND APPLICATIONS OF THE MODEL

Let us begin by examining the least complex situation and proceed in a logical manner to the overall model of the stationary population. Figure 3 illustrates the simplest situation - the study of a group of lives now age x.

Figure 3



One should immediately note the two horizontal axes. The axis labeled "time of life" measures time since the birth of each member of the group, while the axis labeled "time of study" measures time since the initialization of the study. I believe that, in reality, only one axis exists - the time axis. The time at any point (ie. the value assigned to any point) is relative to the point from which one measures time. Stated more simply, the time value assigned to any point along the time axis depends upon the point in time which one defines to be the origin.

Now let us define the areas which are labeled A and B in figure 3 . Area A represents the total past lifetime of a group of lives currently (ie. time of study = 0) age x. Area B represents the total future lifetime of the same group. Of course, the sum of the two areas equals the total lifetime of a group of lives now age x. It should be apparent that the following are true:

$$\text{Area } A = x \cdot l_x$$

$$\text{Area } B = \int_x^{\omega} l_y \, dy, \text{ where y is the "time of life"}$$

$$= \int_0^{\omega - x} l_{x+t} \, dt, \text{ where t is the "time of study"}$$

$$= T_x$$

$$\text{Area } (A+B) = x \cdot l_x + T_x = F_x$$

The integrals used to evaluate area B are quite correct; however, they are also quite deceiving. The integrals shown above give the impression that the future lifetime of a group of l_x lives is obtained by summing verticle lines, as depicted by figure 4a. [This is correct but confusing] Actually, the area is derived by summing horizontal lines, as shown in figure 4b . Each horizontal line represents the future lifetime of those $l_{x+t} \cdot \mu_{x+t} \, dt$ individuals now age x, who die upon attaining age x+t.

(8)

Figure 4a

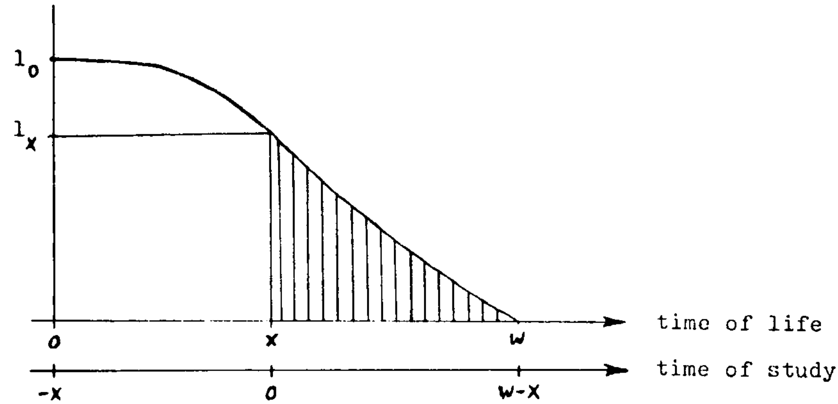
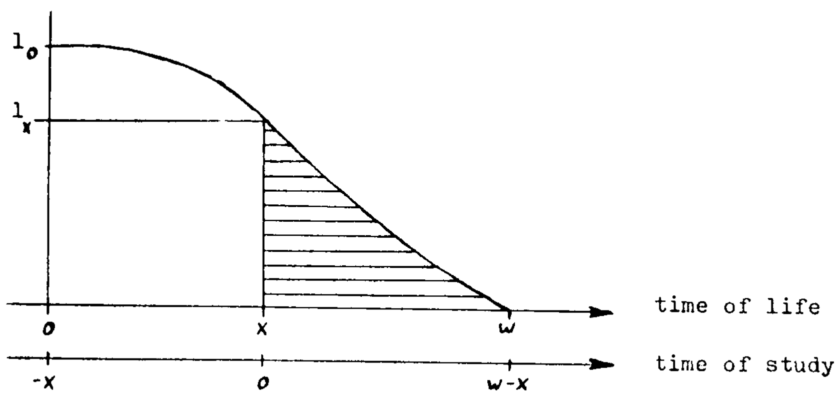


Figure 4b



From the foregoing analysis and by carrying Dr. Nesbitt's viewpoint one step further, it follows that a group of lives now age x is composed of $l_{x+t}\cdot\mu_{x+t}\,dt$ individuals who survive exactly t years ($0 \leq t \leq w-x$). Therefore, the total future lifetime of a group of l_x lives may be expressed as :

$$ {}_0\int^{w-x} t\cdot l_{x+t}\cdot\mu_{x+t}\,dt = $$

$$ (-t\cdot l_{x+t}\,|_0^{w-x} \; - {}_0\int^{w-x} 1\cdot -1\cdot l_{x+t}\,dt = $$

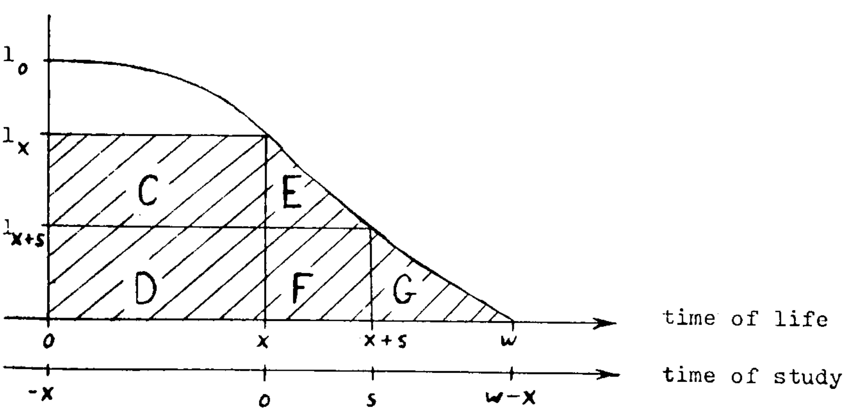
$$ 0 \; + {}_0\int^{w-x} l_{x+t}\,dt = $$

$$ {}_x\int^{w} l_y\,dy $$

Perhaps if a student were introduced to the concept of future lifetime in the above fashion, the student would more readily understand other stationary population problems dealing with future lifetime.

Let us now examine the result of partitioning areas A and B of figure 3. Figure 5 provides us with a model for our study.

Figure 5



Area E and the sum of areas F and G are two areas of major interest. Area E represents the total future lifetime of those lives now age x, who die before attaining age x+s . The sum of areas Fand G represents the total future lifetime of those individuals now age x, who survive to age x+s.[3] Using figure 5, a formula for area E may easily be derived as follows:

(1) Area F = $s \cdot l_{x+s}$

(2) Area G = T_{x+s}

(3) Area (E+F+G) = T_x

(4) Area E = $T_x - T_{x+s} - s \cdot l_{x+s}$

The identical area may be obtained by evaluating $\int_0^s t \cdot l_{x+t} \mu_{x+t} dt$.

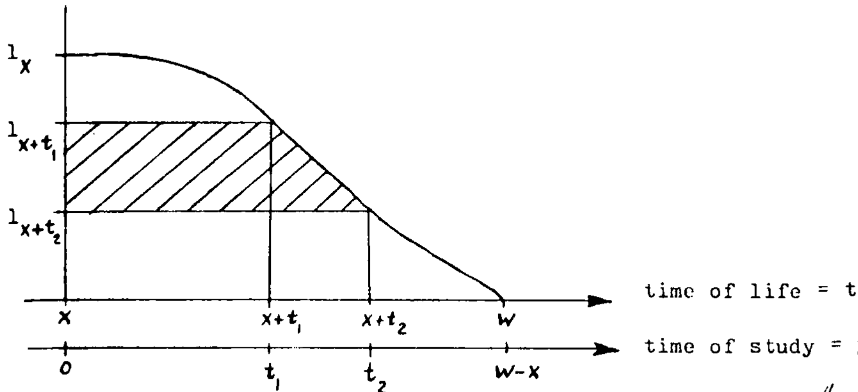
Before studying a group of lives now age x and over, let us consider the problem of finding the future lifetime of those lives now age x, who die between t_1 and t_2 years in the future (between ages $x+t_1$ and $x+t_2$). From figure 6 we can derive an appropriate formula for the shaded region.

$$\begin{aligned} \text{Area} &= (l_{x+t_1} - l_{x+t_2}) \cdot t_1 + T_{x+t_1} - T_{x+t_2} - (t_2 - t_1) \cdot l_{x+t_2} \\ &= T_{x+t_1} - T_{x+t_2} + t_1 \cdot l_{x+t_1} - t_2 \cdot l_{x+t_2} \\ &= (T_{x+t_1} + t_1 \cdot l_{x+t_1}) - (T_{x+t_2} + t_2 \cdot l_{x+t_2}) \\ &= {}_{t_1}H_x - {}_{t_2}H_x, \text{ where } {}_sH_x = T_{x+s} + s \cdot l_{x+s} \end{aligned}$$

The same expression for the area would result if one evaluated $\int_{t_1}^{t_2} t \cdot l_{x+t} \mu_{x+t} \, dt = \int_{x+t_1}^{x+t_2} (y-x) \cdot l_y \mu_y \, dy$.

[3]The significance of this area will become evident shortly in this paper.

Figure 6



$_sH_x = s\, l_{x+s} + T_{x+s}$

The function $_sH_x$ is very useful. The value of $_sH_x$ equals the sum of areas F and G in figure 5, and represents the total future lifetime of those persons now age x, who survive s years (until age x+s). In addition, the function $_sH_x$ may be used to determine the total future lifetime of a group of lives now age x (Area B of figure 3), and the total future lifetime of a group of lives now age x, who die before age x+s (Area E of figure 5).

$$\begin{aligned}\text{Area B} &= {}_0H_x - {}_{w-x}H_x \\ &= (T_{x+0} + 0\cdot l_{x+0}) - (T_w + (w-x)\cdot l_w) \\ &= T_x\end{aligned}$$

$$\begin{aligned}\text{Area E} &= {}_0H_x - {}_sH_x \\ &= T_x - (T_{x+s} + s\cdot l_{x+s}) \\ &= T_x - T_{x+s} - s\cdot l_{x+s}\end{aligned}$$

In the beginning, I believed that the present definition of ${}_{s}H_{x}$ would be adequate; however, further study revealed that this was not so. I found it essential to refine the definition of the function as follows:

$$_{f(x)-x}H_{x} = T_{f(x)} + (f(x) - x)\cdot l_{f(x)}$$

In the new definition, f(x) is a function of the present age of a group of lives and the difference, f(x) - x = t(x) , yields the length of time that elapses before an individual age x attains age f(x). In figure 6 on page 11, the variable t (time of study) would be labeled t(x) and the variable y (time of life) would be labeled f(x) under the new definition. The appropriateness of the new definition will become apparent in the latter portion of this paper. Before continuing, it should be noted that, if f(x) = x+s, then $_{f(x)-x}H_{x} = {}_{s}H_{x} = T_{x+s} + s\cdot l_{x+s}$.

Having a thorough understanding of the study of a group of lives now age x, let us now direct our attention toward the study of a group of lives now age x and over. At the initial time of our study we would find l_y lives who are exactly age y ($x \leq y \leq w$). Thus, above the "ages of entry into the study" axis, we would find an l_x curve as shown in figure 7a. Also, we would discover the identical l_x curve above the time axis, as shown in figure 7b.

Figure 7a

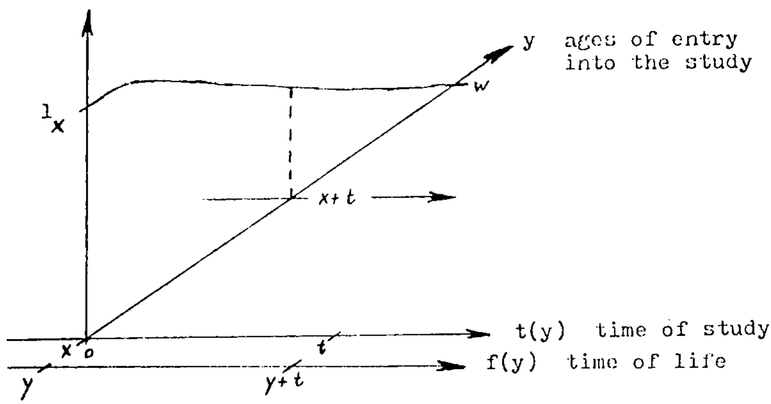
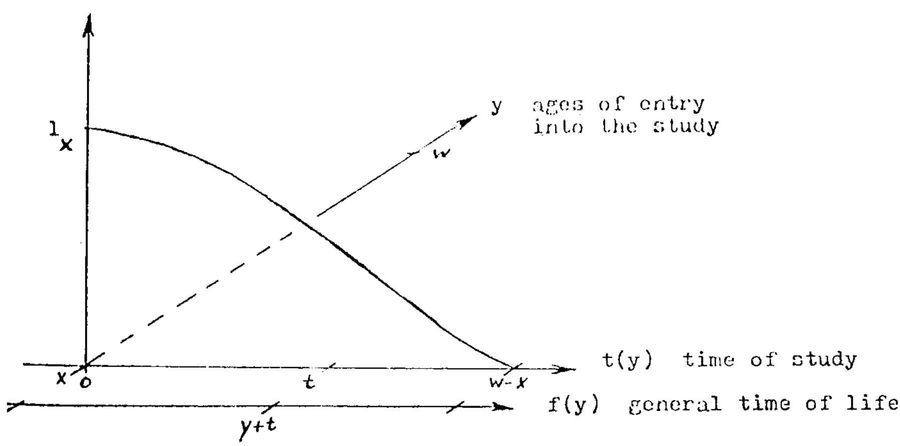


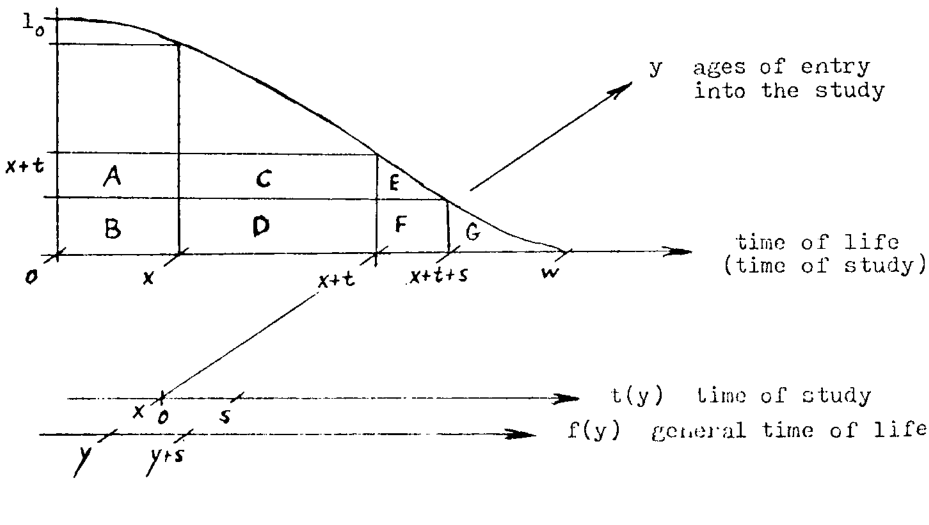
Figure 7b



Figures 7a and 7b illustrate two attributes of a stationary population. First, these figures disclose that the number of lives currently age x and over, equals the total future lifetime of a group of lives now age x. Secondly, they show that it is not necessary to observe a group of lives now age x, until their deaths, in order to procure specific data on the group. In a stationary population, any desired data on a group of l_x lives may be obtained by studying those lives currently age x and over.

In our next step, let us examine a general group of lives now age x+t. Figure 8 provides a model for our study. Any student

Figure 8



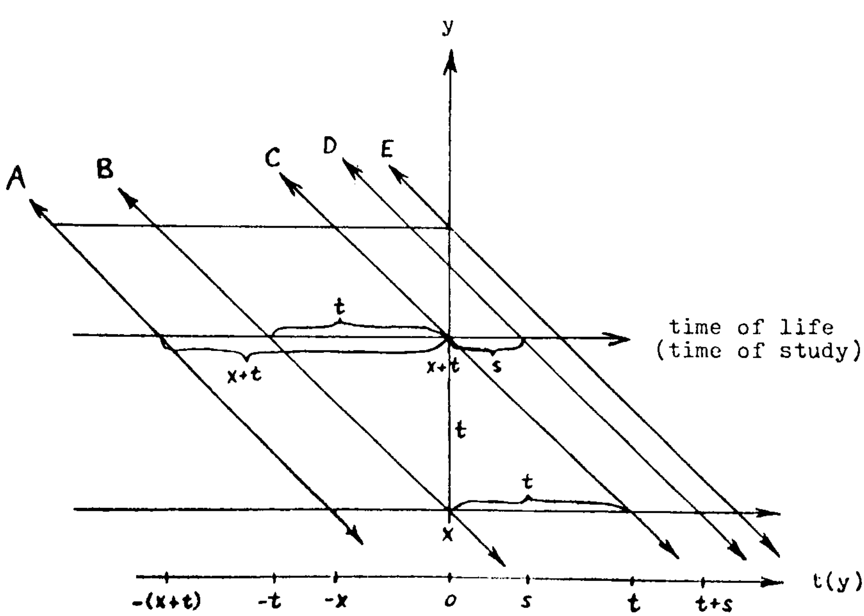
is capable of defining the designated areas; so, I will not define them. In my opinion it is important that the reader notes four characteristics of this general group of l_{x+t} lives:

(1) s years from now, l_{x+t+s} lives age x+t+s will remain
(2) w-(x+t) years from now, there will be $l_w = 0$ survivors.
(3) t years ago, those lives now age x+t were members of a group of lives then age x
(4) x+t years ago, the l_{x+t} lives now age x+t were members of a group of lives age 0 (ie. the l_{x+t} lives now age x+t were born)

Based upon the previous observations, we can note that all of the following statements concerning figure 9 are true:

(1) All lives above line A are age 0
(2) All lives above line B are age x
(3) All lives above line C are age x+t
(4) All lives above line D are age x+t+s
(5) All lives above line E are age w

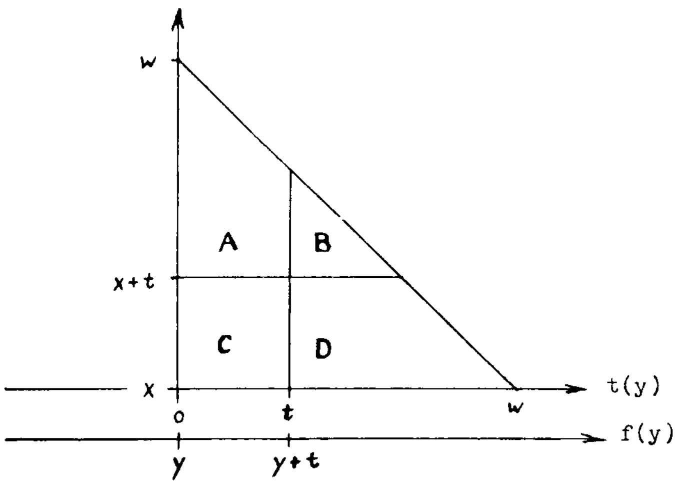
Figure 9



The proof that the past lifetime of a group of lives now age x and over, since attaining age x, is equivalent to the total future lifetime of the same group of lives necessitates the knowledge of the forementioned facts.

Another interesting set of properties which the model possesses may be discovered if one subdivides the total future lifetime of a group of lives age x and over, as shown in figure 10.

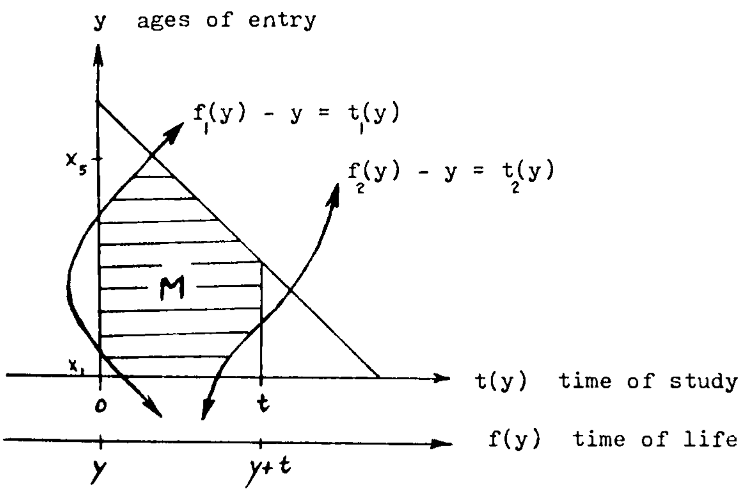
Figure 10



The sum of the area above regions A,B,C, and D is equivalent to Y_x , the total future lifetime of a group of lives now age x and over. Note that the sum of the area above regions A and B, as well as the sum of the area above regions B and D, equals Y_{x+t}. The area above region B is readily seen to be Y_{x+2t}. It follows that the area above both region A and region D equals $Y_{x+t}- Y_{x+2t}$. Finally, the area above region C must then be $Y_x -2\cdot Y_{x+t}+ Y_{x+2t}$.

Before applying the model to a specific problem, let us see how the model may be applied to a general stationary population problem. Let us use the model to determine the total future lifetime of those individuals now between the ages of x_1 and x_5, who die within t years and between ages $f_1(y)$ and $f_2(y)$, where y represents the present age of a typical group of l_y lives. The region over which we must integrate is shown in figure 11.

Figure 11

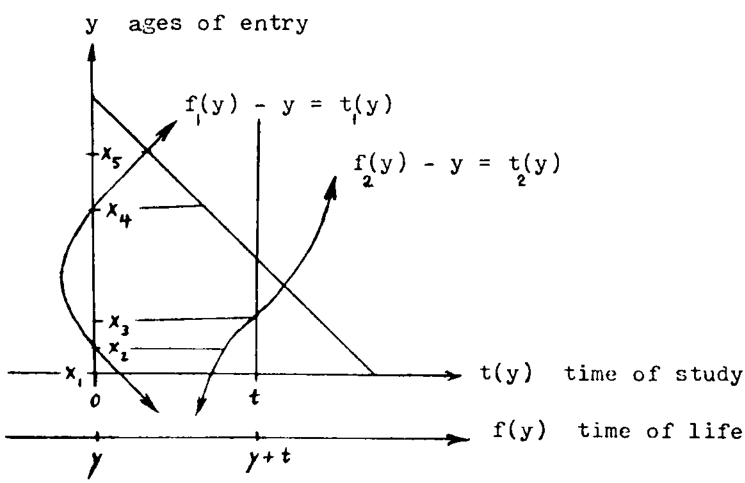


Note that the region does not extend beyond the line t(y) = 0, where t(y) refers to the time of study. This is due to the fact that any person now living (ie. alive at "time of study" = 0) can not possibly die in the past (time of study < 0). Alternatively, one may view this as an age restriction. No individual now age y can die at an age less than y. This restriction, although not specifically mentioned , is found in some stationary population problems and the student should be aware of its existence.

In order to determine the total future lifetime of those who die in region M, region M must be partitioned as depicted by figure 12. The area above region M may be obtained by summing the following integrals:

$$\int_{x_1}^{x_2} \left({}_{f_1(y)-y}H_y - {}_{f_2(y)-y}H_y \right) dy$$
$$+\int_{x_2}^{x_3} \left({}_{f_3(y)-y}H_y - {}_{f_2(y)-y}H_y \right) dy \text{ , where } f_3(y) = y$$
$$+\int_{x_3}^{x_4} \left({}_{f_3(y)-y}H_y - {}_{f_4(y)-y}H_y \right) dy \text{ , where } f_4(y) = y+t$$
$$+\int_{x_4}^{x_5} \left({}_{f_1(y)-y}H_y - {}_{f_4(y)-y}H_y \right) dy$$
$$= \int_{x_1}^{x_2} \left({}_{f_1(y)-y}H_y - {}_{f_2(y)-y}H_y \right) dy$$
$$+\int_{x_2}^{x_3} \left({}_{t_3(y)}H_y - {}_{f_2(y)-y}H_y \right) dy \text{ , where } t_3(y) = 0$$
$$+\int_{x_3}^{x_4} \left({}_{t_3(y)}H_y - {}_{t_4(y)}H_y \right) dy \text{ , where } t_4(y) = t$$
$$+\int_{x_4}^{x_5} \left({}_{f_1(y)-y}H_y - {}_{t_4(y)}H_y \right) dy$$

Figure 12



If $l_{f(y)}$ is substituted for ${}_{f(y)-y}H_y$, the number of lives dying within the region may be found. To derive the total lifetime of those who die, one may either substitute $F_{f(y)} = f(y)\cdot l_{f(y)} + T_{f(y)}$[4] for ${}_{f(y)-y}H_y$ or substitute F_y for l_y and G_y for T_y in the expression obtained for the number of lives dying. In addition, the total future lifetime lived in region M by those individuals now between ages x_1 and x_5, may be found by substituting $T_{f(y)}$ for ${}_{f(y)-y}H_y$.

In the majority of problems, f(y) is a linear function of y. It is instructive to note that, if $f(y) = a\cdot y + b$, the following are true:

(1) $\frac{d}{dy} T_{ay+b} = -a\cdot l_{ay+b}$

(2) $\frac{d}{dy} Y_{ay+b} = -a\cdot T_{ay+b}$

(3) $\frac{d}{dy}\, {}_{(a-1)y+b}H_y = -l_{ay+b}\cdot a - [(a-1)\cdot y + b]\cdot l_{ay+b}\cdot\mu_{ay+b}\cdot a + (a-1)\cdot l_{ay+b}$

(4) $\int l_{ay+b}\, dy = \frac{-1}{a}\cdot T_{ay+b}$

(5) $\int T_{ay+b}\, dy = \frac{-1}{a}\cdot Y_{ay+b}$

(6) $\int {}_{(a-1)y+b}H_y\, dy = \frac{-(2\cdot a - 1)}{a^2}\cdot Y_{ay+b} - \frac{(a\cdot y + b - y)}{a}\cdot T_{ay+b}$

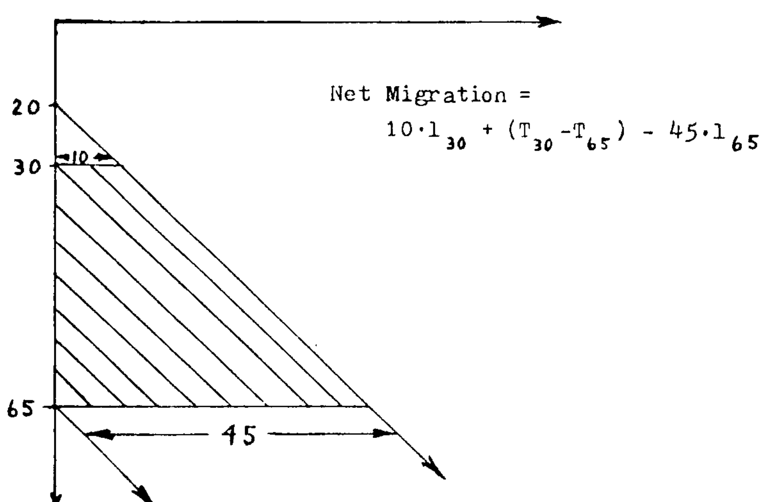
(7) $\int F_{ay+b}\, dy = \frac{-2}{a}\cdot Y_{ay+b} - \frac{(a\cdot y + b)}{a}\cdot T_{ay+b}$

[4] Based upon the Grace-Nesbitt function $F_x = x\cdot l_x + T_x$.

If the coefficient of y in f(y) is an element of the set $\{0,1\}$, the "in-and-out" method may be easily used to derive an expression for the number of people dying (and consequently their total lifetime) within a region whose borders are determined by $t(y) = f(y) - y = (a-1)\cdot y + b$, $a \in \{0,1\}$ and $b \in \mathbb{R}$. If $a \notin \{0,1\}$, the "in-and-out" method may not be applied.

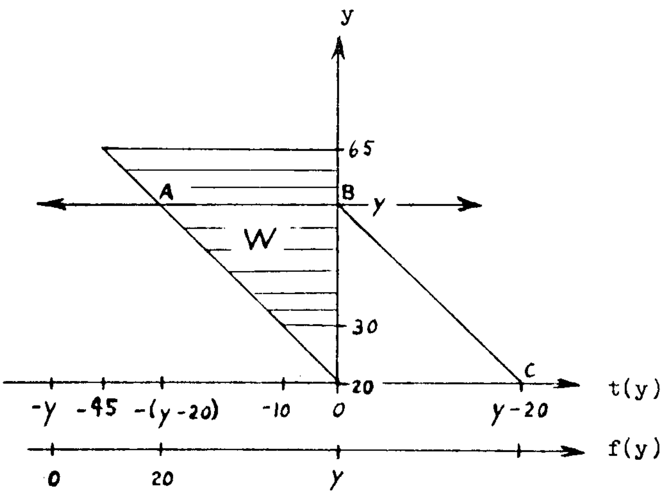
Mr. Veit's "in-and-out" method is a valuable tool if one knows how and when to use it. In Mr. Veit's T.S.A. article, he demonstrated how one could use the method to determine the total past lifetime since age 20, of a group of lives now between the ages of 30 and 65. In his paper, he obtained an expression for the net migration through the shaded area in figure 13 (shown as figure 4 in Mr. Veit's article) and derived the desired result by substituting T_x for l_x and Y_x for T_x.

Figure 13



In my model, the area which represents the total past lifetime since age 20, of the same group of lives, lies above region W in figure 14. If we view a typical age y, we would discover a rectangle of length y-20 and height l_y above region W. The area of the rectangle is, of course, $(y-20)\cdot l_y$.

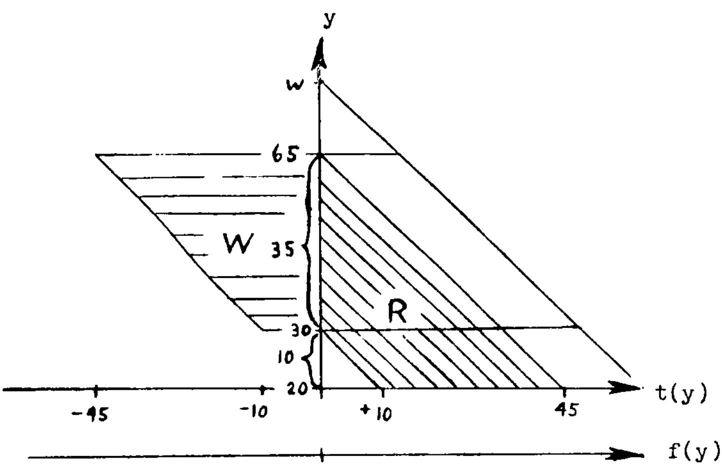
Figure 14



Recall from figure 9, that all lives along a line which passes through the (y,t) plane as line $\overleftrightarrow{BC}$ in figure 14, are age y. If one integrates with respect to time over ages 20 through age y, the sum of the weights applied to each group of l_y lives along line segment $\overline{BC}$ would be y-20 . The resulting area beneath the curve above line segment BC must be $(y-20)\cdot l_y$. Because $(y-20)\cdot l_y$ is equivalent to the past lifetime since age 20 of a group of individuals now age y, it follows that $\overline{BC}$ is a projection of $\overline{AB}$

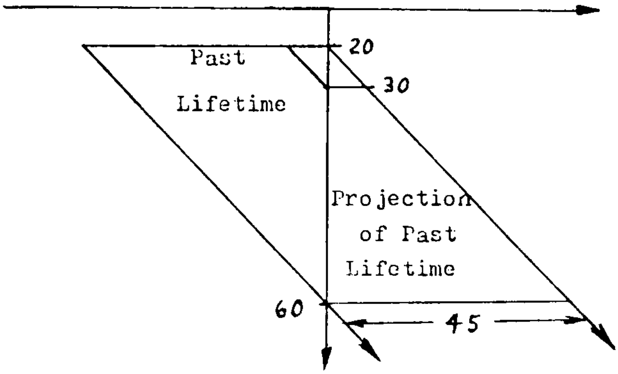
o to future lifetime. If the past lifetime of every age under study is projected onto the future lifetime in a similar manner, we would obtain region R in figure 15. The area above region R equals the are above region W. A formula for the area above region R is $10\cdot(T_{30}-T_{65}) + (Y_{30}-Y_{65} - 35\cdot T_{65})$. Note that the net migration through the region is $10\cdot l_{30}+(T_{30}-T_{65}) - 45\cdot l_{65}$ as Mr. Veit obtained using figure 13.

Figure 15



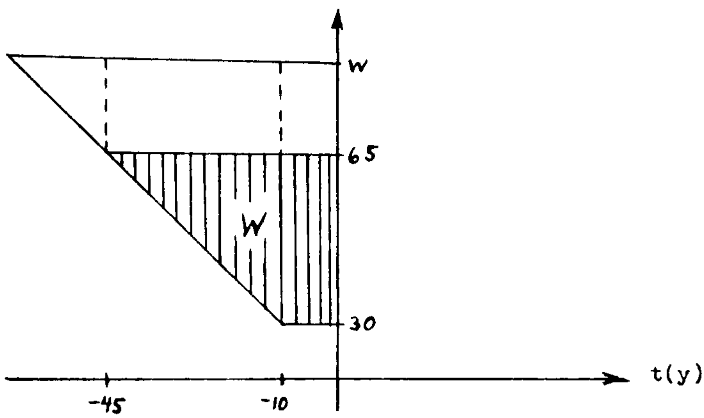
Based upon the preceding discussion, it should now be evident that the area which Mr. Veit utilizes to determine the total past lifetime since age 20, is a projection of the past lifetime onto future lifetime as shown in figure 16.

Figure 16



From our discussion, it may appear that it necessary to project past lifetime onto future lifetime in order to solve problems involving past lifetime. However, the area above region W in figure 14 can be derived by summing the area above the vertical lines shown in figure 17. This alternative approach may prove more valuable to students.

Figure 17



The final portion of this paper is devoted to the derivation, and geometrical verification, of an expression for the total future lifetime of those individuals now between the ages of x and x+t, who die before age x+t. The area which we are attempting to determine is pictured in figure 18a. The area shown in figure 18a lies above the region labeled P in figure 18b.

Figure 18a

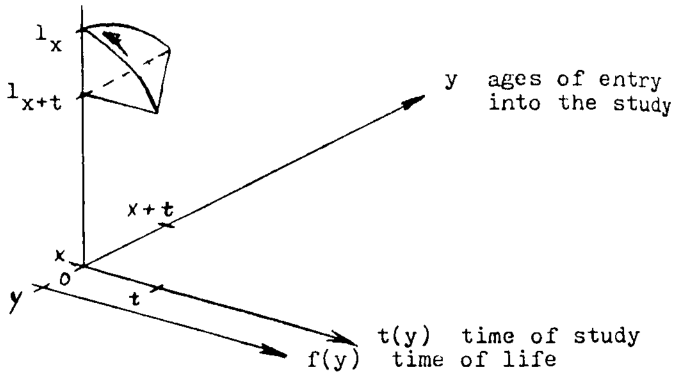
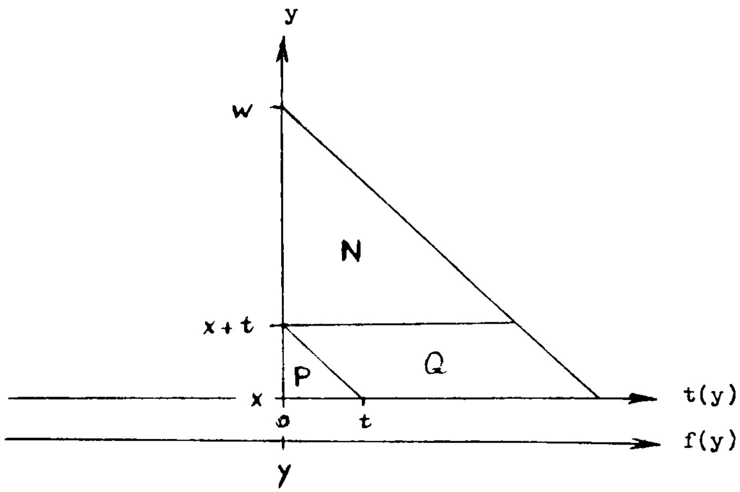


Figure 18b



The number of individuals who die within region P (the net migration through [illegible] region) is

$$\int_{x}^{x+t} (l_y - l_{x+t})\, dy = T_x - T_{x+t} - t \cdot l_{x+t} \, .$$

Coincidentally, the above expression is equivalent to the total future lifetime of those lives now age x who die before age x+t.

The total future lifetime of those who die in region P may be obtained in the following manner:

$$\int_{x}^{x+t} {}_{f_1(y) - y}H_y - {}_{f_2(y) - y}H_y \, dy, \text{ where } f_1(y) = y \text{ and } f_2(y) = x+t$$

$$= \int_{x}^{x+t} ({}_0H_y - {}_{x+t-y}H_y)\, dy$$

$$= \int_{x}^{x+t} (T_{y+0} + 0 \cdot l_{y+0}) - (T_{x+t} + (x+t-y) \cdot l_{x+t})\, dy$$

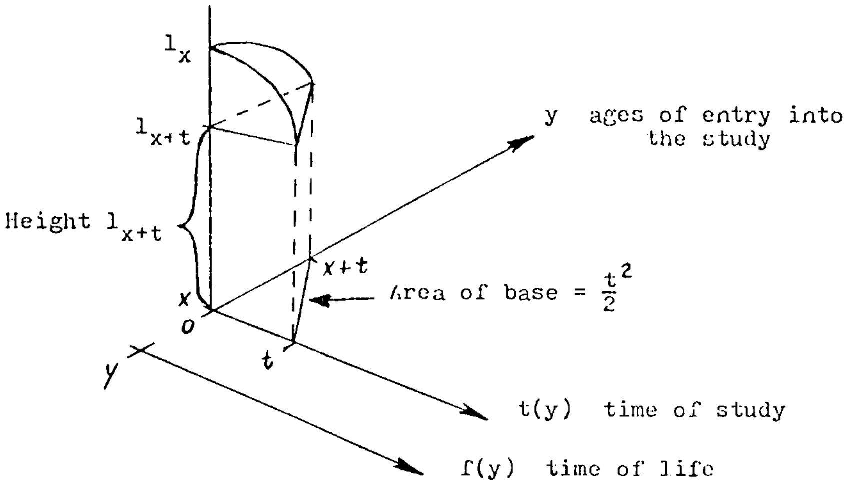
$$= Y_x - Y_{x+t} - t \cdot T_{x+t} - \frac{t^2}{2} \cdot l_{x+t} \, .$$

Now let us confirm that the above formula is valid. The sum of the area above regions N, P, and Q in figure 18b is Y_x. It should be obvious that the area above region N equals Y_{x+t}. The area above region Q is not as obvious. An expression for the area is

$$\int_{x}^{x+t} T_{x+t}\, dy = t \cdot T_{x+t} \text{, where y refers to an "age of entry".}$$

It follows that the total area above region P can be expressed as $Y_x - Y_{x+t} - t \cdot T_{x+t}$. However, we are not searching for an expression for the entire area above region P (shown in figure 19); we wish to determine a formula for the shaded region in figure 18a.

Figure 19



If we subtract the volume of the wedge with height l_{x+t} and with a base of $\frac{t^2}{2}$, we will obtain the volume of the shaded region of figure 18a. Therefore, the total future lifetime of those individuals now between age x and x+t, who die before age x+t is indeed $Y_x - Y_{x+t} - t \cdot T_{x+t} - \frac{t^2}{2} \cdot l_{x+t}$.

BIBLIOGRAPHY

Veit, Kenneth P. , Stationary Population Methods, Transactions of the Society of Actuaries, 1965,XVI, 233-43.

Nesbitt, Cecil J.; Tino, Paulette; et. al. , "Discussion of the Preceding Paper", Transactions of the Society of Actuaries,1965, XVI, 244-64