Canadian Studies in Population, Vol. 23(2),1996, pp. 147-163

Another Look at the Logit Transformation of the Survivorship Function

S. Mitra
Emory University
Atlanta, Georgia

Abstract

In support of the logit transformation of the life table survivorship function *l(x)*, Brass has noted that when a variable assumes values between 0 and 1 its logit varies from $-\infty$ to ∞. Not only that, if the variable is a function of x, which *l(x)* certainly is, then the logit "will vary almost linearly with the x variable." If that is the case, then the logits of the life table functions of any two life table survivorship functions must exhibit strong linear relationship and indeed they do. In this paper it has been shown that while that strong linear relationship between the two logits is an empirical fact, it is definitely not due to the linearity between the logit of *l(x)* and x. The empirical linear relationship between the two logits can certainly hold for other nonlinear functional relationships between the logit and x. The field of such relationships can be considerably narrowed down by introducing necessary restrictions that the *l(x)* function must meet.

Résumé

À l'appui de la transformation logit de la fonction de survie *l(x)* des tables de survie, Brass note que quand une variable assume des valeurs entre 0 et 1, son logit varie de $-\infty$ à ∞. Plus encore, si la variable est une fonction de x, ce qui est certainement le cas de *l(x)*, le logit varie de façon presque linéaire avec la variable x. Si tel est le cas, les logits des fonctions des tables de survie de deux fonctions de survie quelconques doivent présenster une forte relation linéaire, ce qui se vérifie. Le présent article démontre que, bien que la relation linéaire entre les deux logits soit un fait empirique, elle n'est certes pas attribuable à la linéarité entre le logit de *l(x)* et x. La relation linéaire empirique entre les deux logits vaut certainement pour d'autres relations fonctionnelles non linéaires entre le logit et x. Le champ de relations comme celles-ci peut être considérablement réduit par l'introduction de restrictions nécessaires auxquelles la fonction doit être assujettie.

Key words: life table, survivorship function, nonlinear model, logit transformation

Introduction

Application of the logit transformation to the survivorship function *l(x)* proposed by Brass (1975) was a significant step towards statistical modeling of life tables. He recommended the transformation since a) the logit of *1- l(x)* ranges from $-\infty$ to ∞ against a range of 0 to 1 of *1- l(x)* and further b) the logit varies almost linearly with age *x*. Accordingly, he theorized that the logits of two life tables must exhibit a strong linear relationship which has been validated by many examples (Keyfitz, 1991).

The purpose of this paper is to show that although a linear relationship (if it is so) between the logit of *1- l(x)* and *x* guarantees linearity between the logits of any two life tables, it is by no means a necessary condition. In fact, it will be shown that this condition is not only not necessary but it is also not expected to hold for the *l(x)* function. Consequently, the empirical linear relationship between the two logits must derive its support from some type of a nonlinear relationship between the logit of *1- l(x)* and *x*.

An attempt will be made in this paper to develop a justification for the logit transformation of the survivorship function from another perspective. Thereafter, the mathematical functional form describing its relationship with age will be derived by taking into account certain limiting conditions. But first, let us begin by looking into the consequences of a linear relationship between the logit of *1- l(x)* and *x*.

The Logit of *1- l(x)* and its Characteristics

Brass (ibid) has used the following definition of the logit transformation of p $(0 \leq p \leq 1)$ given by

$$logit\, p = \tfrac{1}{2}\, ln\, \frac{p}{1-p} \qquad (1)$$

Thus for $p = 1 - l(x)$, (1) becomes

$$logit\, p(1 - l(x)) = \tfrac{1}{2}\, ln\, (\frac{1 - l(x)}{l(x)}) \qquad (2)$$

It is easy to see that as x increases from 0 to some maximum age a, $l(x)$ declines from 1 to 0 and the logit of $1 - l(x)$ increases from $-\infty$ to ∞. If the logit is linearly related to age, then we should be able to write (disregarding the multiplier of 1/2),

$$ln\left(\frac{1 - l(x)}{l(x)}\right) = a + bx \qquad (3)$$

or

$$l(x) = \frac{1}{1 + e^{a+b}} \qquad (4)$$

It is obvious that if (3) holds then either a or b or both must be nonzero. In that case (4) will not meet conditions like $l(0) = 1$ and $l(\ ) = 0$. The boundary conditions can be met only if for nonzero b, x has a range of $-\infty$ to ∞. Thus, for $b>0$, $l(-\infty) = 1$ and $(\infty) = 0$ and vice versa for $b<0$. Obviously, such limits cannot be applied on the survivorship function.

It can be shown that a distribution function of the form of (4) has a density function like (Brass, ibid)

$$y = \frac{1}{2\beta} sech^2[(Z - \alpha)/\beta] \quad -\infty \leq Z \leq \infty \qquad (5)$$

The curve generated by y is symmetrical and bell shaped. In no way can it be made to resemble the comparable curve in the life table i.e., the curve generated by the age distribution of deaths. Brass (ibid) however, cites (5) as an example of the logit being a linear function of age (his formula in the paper has a typographical error as it begins with 1/2 instead of $1/(2\beta)$).

We conclude this section by pointing out another oddity of the linear relationship. Differentiating both sides of (3) and equating, we get

$$\frac{l(x)\mu(x)}{1 - l(x)} + \mu(x) = b \qquad (6)$$

where

$$-\frac{1}{l(x)}\frac{dl(x)}{dx} = \mu(x) \qquad (7)$$

is the force of mortality. Simplifying (6), we get

$$\mu(x) = b\,(1 - l(x)) \tag{8}$$

which requires that $\mu l(x)$ has to increase with x and it has a minimum value of 0 at $x = 0$ and a maximum value of b at the last age. This is contrary to the well known pattern of the variation of $\mu(x)$.

Therefore, any explanation for the linearity between the logits of the survivorship functions say, $l(x)$ and $l^*(x)$ of any two life tables, described by

$$logit\ (1 - l(x)) = a + b\ logit\ (1 - l^*(x)) \tag{9}$$

should be sought not in (3) but somewhere else. At this point it can be seen that if the logit function can not be expressed as a linear function of age it must be a nonlinear function. If that functional relationship be such that

$$logit\ (1 - l(x)) = u + vn\ (x) \tag{10}$$

and

$$logit\ (1 - l^*(x)) = u^* + v^*n\ (x) \tag{11}$$

where $n(x)$ is a nonlinear function of x which remains the same for all life tables. Then it can be shown that (9) will hold for

$$u = a + bu^* \tag{12}$$

and

$$v = bv^* \tag{13}$$

Our next problem would then have been to search for a functional form of this $n(x)$ if it had not been recently shown that this simple linear model (9) fails to meet a boundary condition unless $b=1$ (Mitra, 1995). Forcing that restriction, however, lowers the quality of the fit. Therefore, acknowledging the advantages of the logit transformation we now turn our attention to look for a possible underlying relationship between $l(x)$ and x.

Another Justification of the Logit Transformation

It is obvious that a linear relationship cannot exist between $l(x)$ and $l^*(x)$ unless they are equal for all x. In all other cases $l^*(x)$ plotted against $l(x)$ will generate a curve that will be either concave upward or downward accordingly, as the life expectancy of the former is less or greater than the same of the latter. The simple relationship between the logits of two life tables is, however, possible if among other things the rate of change of the survivorship function $l(x)$ at any age x, i.e., $d\, l(x)/dx$ turns out to be proportional to its value at that age, i.e., $l(x)$ and also to its complement, i.e., $1- l(x)$. This seems quite logical because the pattern of change in $l(x)$ at age x can be expected to be determined by the number of those who have managed to survive to that age and also by those who have not. Such a relationship can be expressed in the form of an elementary differential equation as

$$\frac{dl(x)}{dx} = -k(x)l(x)(1-l(x)) \qquad (14)$$

where $k(x)$ accounts for all other unknown patterns of variation in $l(x)$. The solution of (14) turns out to be expressible in terms of the logit transformation. This may be seen by writing (14) as

$$-\frac{dl(x)}{l(x)(1-l(x))} = -k(x)dx \qquad (15)$$

Integrating both sides of (15), we can write

$$ln\,(\frac{1-l(x)}{l(x)}) = f(x) \qquad (16)$$

where $f(x)$ remains to be determined.

The Search for a Form of $f(x)$

In order to propose a functional form of $f(x)$ we begin by rewriting (14) as (see (7),

$$\mu(x) = f'(x)(1- l(x)) \qquad (17)$$

where $f'(x) = -k(x)$. We know (see discussion following equation 8) that $f'(x)$ cannot be a constant, or that $f(x)$ cannot be a linear function of x. Thus $f(x)$ must be nonlinear function of x, a generalized form of which can have two or more additive components such that we can write

$$f(x) = g(x) + h(x) + \ \ldots \qquad (18)$$

Accordingly, restricting (18) to first two terms (17) can be expressed as

$$\mu(x) = g'(x)(1 - l(x)) + h'(x)(1 - l(x)) \quad (19)$$

As $x \to 0$, let us suppose that

$$\lim_{x \to 0} g'(x)(1 - l(x)) = 0 \quad (20)$$

which will happen when $g'(0)$ is finite since $l(0)=1$. In that event, the second component must have a limiting value (see (16)), like

$$\lim_{x \to 0} h'(x)(1 - l(x)) = 0 \quad (21)$$

It can be seen that (21) holds when

$$h(x) = \ln x \quad (22)$$

or

$$h'(x) = \frac{1}{x} \quad (23)$$

because substituting (23) in (21) we get

$$\lim_{x \to 0} \frac{1 - l(x)}{x} = \lim_{x \to 0} l(x)\mu(x) = \mu(0) \quad (24)$$

by the application of L'Hospital's rule and the fact that $l(0)=1$. Another simple form of $h(x)$ like

$$h(x) = \ln(e^x - 11) \quad (25)$$

or

$$h'(x) = \frac{e^x}{e^x - 1} = 1 + \frac{1}{e^x - 1} \quad (26)$$

also satisfies (21). Since either of these two forms of $h(x)$ meets the boundary condition, we have found good reasons to choose the latter. From theoretical considerations, the function $h(x)$ should preferably have a wider range of variation. A comparison of (22) and (25) reveals that the rate of change of the former is considerably less than that of the latter, and especially so, for large values of x. Next, we have found that the reproducibility of the model is much better with (25) than with (22) as measured by the R^2 coefficient between the model and the actual values (of the order of .99 versus .92). Thus, our choice for $h(x)$ has been based on both theoretical and empirical grounds.

For the other component $g(x)$ of $f(x)$, we may begin with the simple linear form $a + bx$. Note that the nonlinearity of $f(x)$ has, at least in part, been taken care of by $h(x)$ which also fulfills the requirement that the logit function equals $-\infty$ when $x=0$. This linear form of $g(x)$ together with $h(x)$ make $f(x)$ quite large at the last age α which is some age above 100. However, the theoretical value of the logit at that age is ∞ and therefore a case can perhaps be made in support of modifying $g(x)$ or by adding a third component $k(x)$ to (18) in such a way that the condition at the upper extreme can also be met. As a trial solution to this problem, we have chosen to experiment with a simple function like $1/(\alpha - x)$ although there are many others like $-ln(\alpha - x)$, $(e^{\alpha - x} - 1)^{-1}$ etc., fulfilling that requirement.

Accordingly, we can express $g(x)$ as

$$g(x) = a + bx + \frac{c}{\alpha + x} \tag{27}$$

The Graduation Model and its Application

What follows next is the formulation of a mathematical model for the logit function a complete form of which may now be derived from the preceding as

$$ln\left(\frac{1 - l(x)}{l(x)}\right) = ln(e^x - 1) + a + bx + \frac{c}{\alpha - x} \tag{28}$$

Observe that we have not provided for a coefficientof the first term of the right hand side of (22). If we did and that coefficient in any given example turned out to be other than one, the limiting condition (21) would have been violated. Therefore, we have chosen to rewrite (28) as

$$ln\left(\frac{1 - l(x)}{l(x)}\right) - ln(e^x - 1) = a + bx + \frac{c}{\alpha - x} \tag{29}$$

such that the entire left hand side of (29) can be taken as the dependent variable for estimating the parameters a, b and c by the OLS regression. For that exercise the value of α has been arbitrarily set between say 100 and 125 in this study.

Note that the parameters a, b and c thus estimated will produce the minimum value of the sum of squares of deviation for the model equation (29). However, it may not minimize the sum of squares of differences between the actual and the model values of $l(x)$ where the latter derivable from (29) is given by the nonlinear expression

$$l(x) = \frac{1}{1 + (e^x - 1)e^{a+bx+c/(\alpha-x)}} \qquad (30)$$

Accordingly, we have used the parametric values generated by (28) as seed values to produce new estimates for the nonlinear model equation (28).

Results

In Table 1, a goodness of fit index of this nonlinear model mesured by the R^2 values for different values of α starting at 100 and ending at 125 at 5 year intervals has been presented.

Nine life tables were selected for this study covering a range of life expectancy of 53 to 70 years for males and 52 to 78 years for females. Although R^2 values are quite large (over .99) for all values of α, we have arbitrarily decided to set its value at 120 for subsequent investigation. Next for that value of α, the estimated values of the parameters may be seen in Table 2.

It is interesting to note that for all life tables, the parameter b assumes values very close to -1 and has a tendency to increase with life expectancy. However, the three parameters tend to vary simultaneously with changes in levels of mortality. As such, when one of the parameters is allowed to increase, holding the other two constant, it can be seen from (30) that the denominator becomes uniformly larger or $l(x)$ uniformly smaller. Like b, the parameter a is also negative and appears to have a range of variation from -6 to -4 in round numbers for both males and females. The parameter c on the other hand seems to be most sensitive as its range of variation extends from as low as 140 to as high as 360 for the males. The range (185/270) is considerably narrower for the females.

In Table 3, the values of $l(x)$ generated by (30) for three of the nine countries have been presented together with their actual values. The two sets are very similar to one another as one may anticipate from the larger R^2 values. A graphical comparison of the model and the actual values can be made for the same countries from Figures 1 and 2 for the males and the females respectively.

In general, the match between the two sets of figures seem to be quite good. The only criticism of any significance that can be levelled is that for the males of

Table 1: R-Squared Values between Model and Actual Survivorship Function for Different Values of Life Span

Country	Year	Sex	e_0	R-squared					
				100	105	110	115	120	125
Tunisia	1968-69	m	52.71	0.99823	0.99916	0.99938	0.99923	0.99890	0.99848
		f	52.45	0.99960	0.99938	0.99868	0.99779	0.99684	0.99590
Cuba	1969-71	m	53.83	0.99915	0.99978	0.99990	0.99809	0.99957	0.99930
		f	55.63	0.99892	0.99946	0.99945	0.99851	0.99882	0.99841
Columbia	1963-65	m	56.93	0.99778	0.99890	0.99943	0.99965	0.99970	0.99965
		f	59.69	0.99732	0.99876	0.99947	0.99981	0.99993	0.99993
Mauritius	1971-73	m	60.73	0.99554	0.99702	0.99798	0.99861	0.99904	0.99933
		f	65.65	0.99676	0.99846	0.99931	0.99971	0.99987	0.99988
Guyana	1969-71	m	61.98	0.99703	0.99802	0.99863	0.99904	0.99932	0.99952
		f	66.60	0.99689	0.99799	0.99866	0.99910	0.99938	0.99958
Argentina	1969-70	m	62.08	0.99587	0.99725	0.99812	0.99869	0.99907	0.99934
		f	69.65	0.99568	0.99767	0.99875	0.99935	0.99967	0.99982
Guatemala	1972-73	m	70.28	0.99365	0.99588	0.99723	0.99980	0.99864	0.99900
		f	73.68	0.99463	0.99657	0.99775	0.99919	0.99900	0.99932
Martinique	1981-83	m	70.98	0.99841	0.99872	0.99893	0.99909	0.99921	0.99931
		f	75.36	0.99869	0.99910	0.99929	0.99936	0.99935	0.99931
Puerto Rico	1979-81	m	70.24	0.99912	0.99946	0.99964	0.99973	0.99979	0.99981
		f	77.58	0.99740	0.99808	0.99857	0.99892	0.99917	0.99936

Table 2: Values of Model Parameters for Life Span of 120 Years

Country	Male					Female				
	e_0	R-squared	a	b	c	e_0	R-squared	a	b	c
Tunisia	52.71	0.99890	-3.62	-1.02	278.34	52.45	0.99684	-3.48	-1.02	267.38
Cuba	53.83	0.99957	-3.31	-1.00	201.60	55.63	0.99882	-3.37	-1.01	202.21
Columbia	56.93	0.99970	-3.95	-1.01	238.35	59.69	0.99993	-3.95	-1.01	229.05
Mauritius	60.73	0.99904	-5.40	-1.02	359.59	65.65	0.99987	-4.66	-1.01	262.38
Guyana	61.98	0.99932	-5.08	-1.00	248.96	66.60	0.99938	-5.04	-1.00	219.30
Argentina	62.08	0.99907	-5.00	-1.01	277.44	69.65	0.99967	-4.70	-1.01	235.17
Guatemala	70.28	0.99864	-5.04	-1.01	241.39	73.68	0.99900	-5.04	-1.00	210.57
Martinique	70.98	0.99921	-5.80	-0.96	142.14	75.36	0.99935	-5.89	-0.99	216.26
Puerto Rico	70.24	0.99979	-5.33	-0.97	151.47	77.58	0.99917	-6.15	-0.98	185.98

Table 3: Observed and Model Values of Survivorship Function of Three Countries for Life Span of 120 Years

Age	Columbia				Mauritius				Puerto Rico			
	Male		Female		Male		Female		Male		Female	
	e_0 = 56.93		59.69		60.73		65.65		70.24		77.58	
	observed	expected	observed	expected	observed	expected	observed	expected	observed	expected	observed	expected
0	1.00000	1.00000	1.00000	1.00000	1.00000	1.00000	1.00000	1.00000	1.00000	1.00000	1.00000	1.00000
1	0.91850	0.91784	0.92600	0.92330	0.93496	0.94562	0.94648	0.94921	0.97813	0.98885	0.98358	0.99351
5	0.87671	0.87198	0.88155	0.88072	0.91391	0.91566	0.92228	0.92063	0.97559	0.97984	0.98173	0.98855
10	0.86495	0.86425	0.87104	0.87431	0.90972	0.91240	0.91749	0.91721	0.97413	0.97555	0.98055	0.98656
15	0.85806	0.85586	0.86626	0.86739	0.90545	0.90846	0.91465	0.91339	0.97213	0.97039	0.97947	0.98423
20	0.84739	0.84570	0.85892	0.85895	0.89999	0.90301	0.90941	0.90849	0.96588	0.96396	0.97761	0.98136
25	0.83061	0.83330	0.84910	0.84859	0.89380	0.89560	0.90198	0.90218	0.95631	0.95588	0.97532	0.97776
30	0.81212	0.81806	0.83603	0.83579	0.88623	0.88558	0.89202	0.89404	0.94471	0.94563	0.97250	0.97322
35	0.79246	0.79914	0.81947	0.81982	0.87583	0.87200	0.88072	0.88347	0.93167	0.93251	0.96842	0.96737
40	0.77058	0.77541	0.79883	0.79968	0.86076	0.85343	0.86704	0.86958	0.91619	0.91556	0.96262	0.95971
45	0.74405	0.74534	0.77326	0.77400	0.83785	0.82771	0.85063	0.85108	0.89593	0.89343	0.95428	0.94949
50	0.70985	0.70690	0.74180	0.74087	0.80184	0.79156	0.82733	0.82601	0.86760	0.86424	0.94158	0.93552
55	0.66375	0.65743	0.70028	0.69770	0.74726	0.74012	0.79456	0.79146	0.82787	0.82541	0.92204	0.91594
60	0.60008	0.59381	0.64339	0.64114	0.66492	0.66663	0.74576	0.74305	0.77326	0.77350	0.89163	0.88775
65	0.51532	0.51301	0.56778	0.56731	0.55322	0.56354	0.67725	0.67454	0.70197	0.70425	0.84725	0.84609
70	0.41039	0.41389	0.47063	0.47303	0.41459	0.42779	0.57718	0.57839	0.60974	0.61330	0.77975	0.78316
75	0.29424	0.30067	0.35603	0.35907	0.26948	0.27238	0.44525	0.44957	0.49819	0.49838	0.68252	0.68761
80	0.18552	0.18652	0.23498	0.23540	0.14634	0.13307	0.29317	0.29592	0.36656	0.36395	0.54383	0.54746
85	0.09782	0.09191	0.12781	0.12363	0.06493	0.04518	0.15547	0.14939	0.22596	0.22613	0.36996	0.36453

Figure 1. Observed and Expected Values of Survivorship Function for the Males of Three Countries

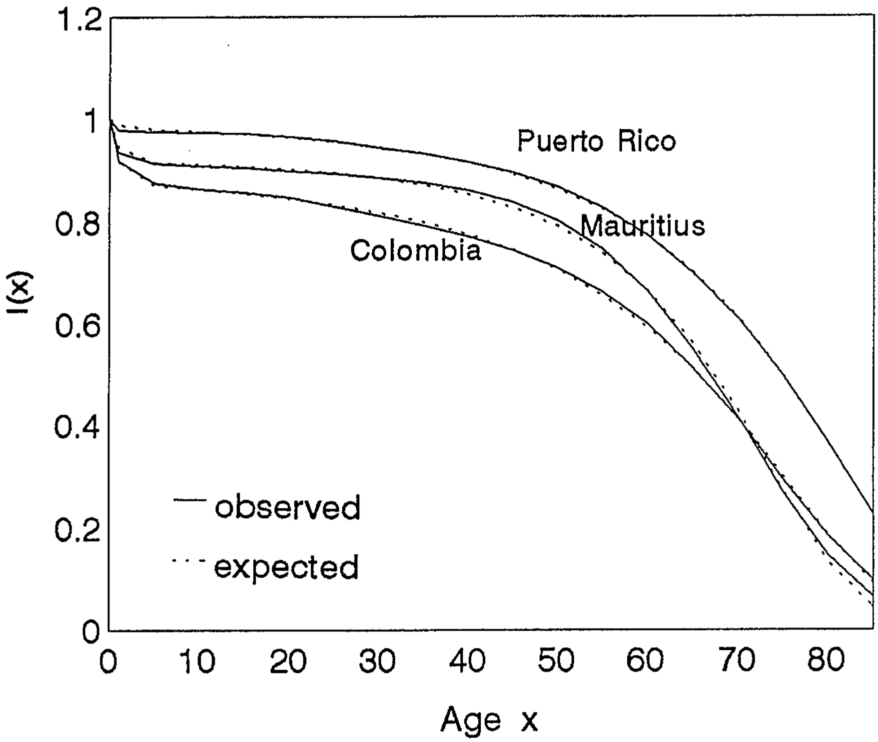
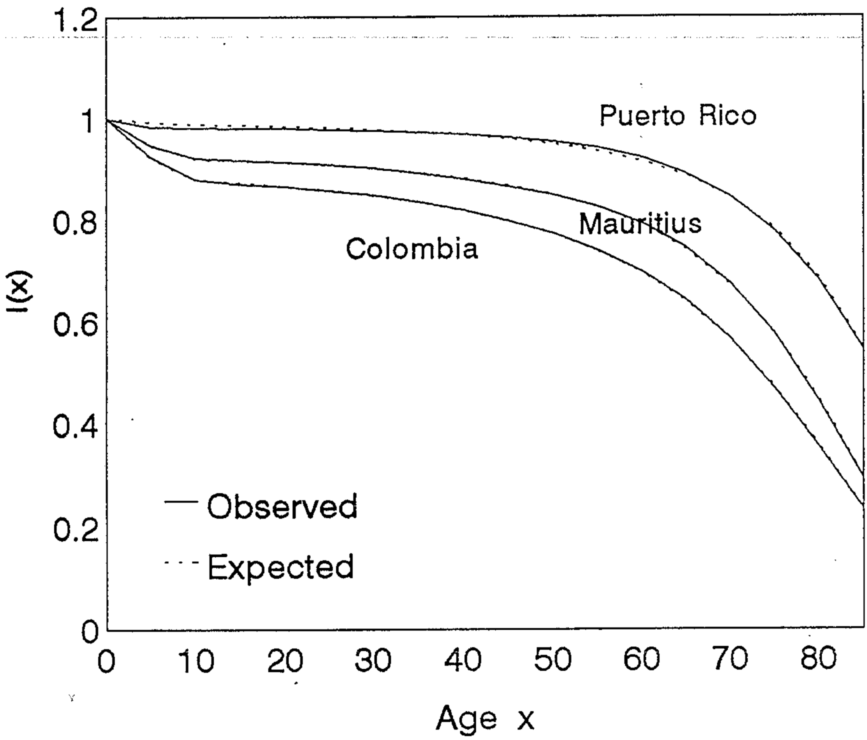


Figure 2. Observed and Expected Values of Survivorship Function for the Females of Three Countries



Mauritius, the decline in *l(x)* at advanced ages is slightly more rapid for the model table. Next, observe that life tables of differentcountries (Mauritius and Colombia males in this illustration) may show crossovers and the three parameter model seems to be flexible enough to adjust its parametric values accordingly.

It is also of some interest to mention that each of the three parameters can vary independently of the other two. Among the nine countries, the largest value of *c* is found in the life table of Mauritius males which has medium but not the smallest life expectancy. Its *a* on the other hand is smaller than that of Puerto Rico which has the highest life expectancy. In order to clarify this point Figure 3 has been drawn by varying one parameter at a time in the model equation of the Columbian males.

Observe that the parameter *b* has almost no effect at the early childhood ages but in spite of its small range of variation, it can affect the mortality curve significantly at other ages for a slight change in its value. The parameters *a* and *c* on the other hand, show significant effect at all ages. By taking partial derivatives it can be seen that the contribution of *a* is more than that of *c* except at the very end of the life span. It may be recalled that the last component with the parameter *c* was introduced in the model primarily to meet the boundary condition at the upper end of the age range.

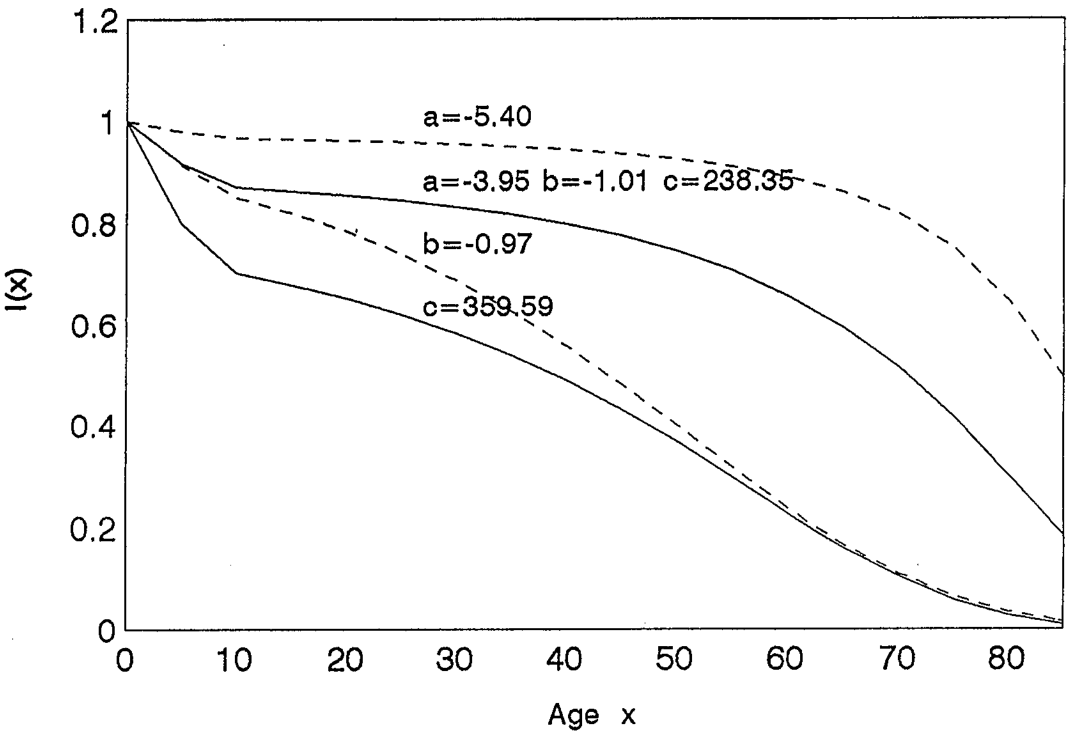
It is interesting to see that the largest value of *b* and *c* found in this set of countries, produced similar effects at older ages (65+) in terms of smaller survivorship probabilities. In this example, it may also be seen that the probabilities can be increased at all ages by simply lowering the value of the parameter *a*. What is apparent from all these is that appropriate combinations of the three parameters should reproduce any distribution of the survivorship probabilities *l(x)* reasonably well. Further investigations are needed to link these parameters with specific patterns of mortality.

Conclusions

The long search for a mathematical function that can adequately describe the pattern of variation of the survivorship probabilities is not over yet. The complexity of the problem derives from nonlinearity of the rate of change of these probabilities. In that endeavor, even the modern high speed computers with programs like TableCurve can not rise to the task since alternatives to a linear function are infinite in numbers. In order to find such a function, if there is one, one has no choice but to postulate (as others have done in the past) the pattern of variation of one or more of the life table functions. Recently, such an attempt by Heligman and Pollard (1980) on the probabilities of dying produced an eight parameter model. Another approach based on the force of mortality (Mitra, 1983

Another Look at the Logit Transformation of the Survivorship Function

Figure 3. Curves of l(x) Drawn by Varying the Parameters One at a Time for Colombia Males



& 1984) generated a three parameter model of the survivorship probabilities *l(x)* which could be linearized by applying the transformation *ln(- ln l(x))*. It has been shown in this paper that an alternative approval, simple and plausible, can also be tried to describe the pattern of variation of *l(x)*. This endeavor resulted in a three parameter nonlinear model which fortunately could be linearized for parametric estimation (equations 28 & 29).

The result seems to be quite encouraging, perhaps more so than the double log model mentioned earlier. The goodness of fit measured by the R^2 coefficient range from .9968 to .9998 for a set of eighteen life tables for nine countries (separately for males and females), covering a wide range of life expectancy (53, 78). Further testing of the model is needed, although there seems to be enough evidence to believe that the model can be used to reproduce any of the currently available distributions of the survivorship function to a reasonable degree of accuracy.

Acknowledgements

This is a revised version of a paper presented at the 1995 meetings of the Southern Demographic Association, Richmond, Virginia (October 19-21). I sincerely appreciate the assistance of Xiahong Mao and Clark Denny, graduate students in our department, in generating the figures and the tables of this paper.

References

Brass, W. 1975. *Methods for Estimating Fertility and Mortality from Limited and Defective Data.* Chapel Hill, NC: Laboratories of Population Sstatistics, An Ocasional Publication.

Heligman, L. and J.H. Pollard. 1980. Age Pattern of Mortality. *Journal of the Institute of Actuaries, Part 1,* 107(434):49-80.

Keyfitz, N. 1991. Experiments in the Projection of Mortality. *Canadian Studies in Population,* 18(2):1-17.

Mitra, S. 1983. Modelling Life Table Functions. *Demography India,* 12(1):115-121.

_____. 1984. On the Characteristics of the Parameters of a Life Table Function. *Genus,* 40(1-2):47-56.

_____. 1995. Boundary Condition Sets Limits of Brass's Model Life Table Parameters. *Canadian Studies in Population,* 22(1):67-78.

Received, January 1996; revised, November 1997