S.Mitra

# The Probability of One Spouse Outliving the Other by Sex and Age

## The Problem

THE probability measure of the event that one of the spouses will survive a given interval of time while the other will not, depends *(a)* on their ages at the beginning of the interval as well as *(b)* on the associated probabilities of surviving and dying in the interval by the respective spouses. The problem is a particular case of the more general example dealing with multiple decrement tables (Jordan, 1967). However, this special case has not received much attention since Depoid (1938) studied the probabilities of a marriage being terminated by the death of the husband or of the wife after a given number of years of marriage. He also mentioned about the eventual probabilities of becoming a widow or a widower.

It may be mentioned that the eventual probability of becoming a widow or a widower can indeed be regarded as an asymptote that is reached as the interval over which such a probability is calculated, is gradually enlarged. In this paper, however, an attempt has been made to directly derive the eventual probabilities of any one of the spouses outliving the other for specific age combinations, that apply either at the time of marriage or at any time thereafter. Needless to say, the dissolution of marriage through separation or divorce is not relevant for the problem defined in this manner.

## Derivation of the Probability Function

For reasons of operational simplicity in the derivation of these results, the survivorship probabilities of the two sexes are regarded as independent of one another. These probabilities are usually obtained from the respective life tables, so that mortality differentials by marital statuses, if any, are also ignored. Under these conditions, the probability that both of the spouses will survive a given interval of time can be obtained as the product of the survivorship probabilities of the two sexes corresponding to their respective ages and the length of the interval. The probabilities of one or both of them dying in that interval can also be easily obtained.

Relaxing the restriction of a specific interval of time, let the probability that a husband $a$ years old will outlive his wife aged $b$, be denoted by $P(a, b)$. Disregarding the possibility of their dying at the same instant of time, the complementary probability, namely

$$Q(a, b) = 1 - P(a, b) \tag{1}$$

provides the probability measure of the same wife's outliving her husband. The probability of their jointly surviving a period of $x$ years can be expressed as

$$C(a + x, b + x) = \frac{l_m(a + x)\, l_f(b + x)}{l_m(a)\, l_f(b)}\,, \tag{2}$$

in which $l_m(a + x)$ and $l_f(b + x)$ are the male and the female probabilities of survival from birth to ages $a + x$ and $b + x$ respectively which coincides with the end of the time interval while $l_m(a)$ and $l_f(b)$ are the corresponding probabiities at the beginning of the interval.

Next, the probability that the female spouse will die at age $b + x$, leaving her male partner a widower, can be written as

$$C(a + x, b + x)\, \mu_f(b + x)\, dx \tag{3}$$

in which $\mu_f(b + x)$ is the force of mortality at age $b + x$ for the females, that is,

$$\mu_f(b + x) = -\frac{d}{dx}\, l_f(b + x) / l_f(b + x)\,. \tag{4}$$

Therefore, the eventual probability of the male's outliving the female spouse can be obtained from

$$P(a, b) = \int_0^{\alpha(a, b)} C(a + x)\, \mu_f(b + x)\, dx, \tag{5}$$

where $\alpha(a, b)$, the upper limit of the integral, depends on the values of $a$, $b$ as well as on the life spans of the two sexes. For all practical purposes, however, the values of $\alpha(a, b)$ may be left unspecified.

Because of (2), $P(a, b)$ can be alternatively expressed as

$$P(a, b) = \frac{\int_0^{\alpha(a, b)} l_m(a + x)\, l_f(b + x)\, \mu_f(b + x)\, dx}{l_m(a)\, l_f(b)} \tag{6}$$

and the complementary probability as

$$Q(a, b) = \frac{\int_0^{\alpha(a, b)} l_m(a + x)\, l_f(b + x)\, \mu_m(a + x)\, dx}{l_m(a)\, l_f(b)}. \tag{7}$$

It is easy to see that (6) and (7) satisfy (1) as they should. This is because

$$-\frac{d}{dx}\left[l_m(a + x)\, l_f(b + x)\right] = l_m(a + x)\, l_f(b + x)\left[\mu_m(a + x) + \mu_f(b + x)\right] \tag{8}$$

and thus the sum of the integrals in the numerators of (6) and (7) simplifies to $l_m(a)\, l_f(b)$, their common denominator.

Another expression for $P(a, b)$ may be derived by noting the equality

$$P(a, b) = \frac{P(a, b)}{P(a, b) + Q(a, b)} \tag{9}$$

so that, a combination of (6) and (7) results in the expression

$$P(a, b) = \frac{\int_0^{\alpha(a, b)} l_m(a + x)\, l_f(b + x)\, \mu_f(b + x)\, dx}{\int_0^{\alpha(a, b)} l_m(a + x)\, l_f(b + x)\left[\mu_m(a + x) + \mu_f(b + x)\right] dx}. \tag{10}$$

## Approximate Algebraic Solution of $P(a, b)$

What follows next is the description of a method suggested for the reduction of the ratio of the integrals in (10). First, it is acknowledged that, in general, the force of mortality can be regarded as a reasonably smooth and a monotonically increasing function of age in the age interval that excludes the childhood years. Also known is the fact that in such an interval, the function can very well be approximated by the Gompertz curve, namely,

$$\mu(x) = BC^{x}. \tag{11}$$

The limitations of the model, found for the most part in the old ages, are known to be true primarily with respect to mortality experiences observed during an interval of time, rather than those that are applicable to generation mortality (Spiegelman, 1969). Consequently, the model can be expected to provide a fair approximation of the mortality experiences, especially in those countries, where the patterns of mortality exhibit little or only minor changes over time.

It is also known that the force of mortality is not affected much by variations in $C$ at the younger ages, and therefore it is generally approximated by an average of its values observed at higher ages. Table 1 shows the values of $C$ for

TABLE 1—ESTIMATED VALUES OF $C$ OF THE GOMPERTZ MODEL $\mu(x) = BC^{x}$ FOR U.S. BY SEX FOR THE YEAR 1973 BEGINNING AGE 30

| *Age* | *30* | *35* | *40* | *45* | *50* | *55* | *60* | *65* | *70* | *75* |
|---|---|---|---|---|---|---|---|---|---|---|
| $C$ (male) | 1.06 | 1.08 | 1.10 | 1.09 | 1.10 | 1.09 | 1.08 | 1.08 | 1.09 | 1.07 |
| $C$ (female) | 1.09 | 1.09 | 1.09 | 1.08 | 1.09 | 1.08 | 1.08 | 1.11 | 1.11 | 1.09 |

SOURCE: 1973 U. S. Life Tables.

the two sexes, obtained from successive five year age intervals ($n = 5$) beginning from age 30, for the 1973 U. S. Life Tables (Vital Statistics of the U.S. 1973, DHEW), as

$$C^{n} = \frac{\int_{0}^{n} \mu(y + n + x)\,dx}{\int_{0}^{n} \mu(y + x)\,dx} = \frac{ln\,[l\,(y + n)/l\,(y + 2n)]}{ln\,[l\,(y)/l\,(y + n)]}. \tag{12}$$

Interestingly enough, the parameter $C$ shows only minor variations by age and sex. This is quite logical in the sense that, under normal conditions, the patterns of mortality of the two sexes cannot be unrelated with one another. For all practical purposes therefore, $C$ can be assumed as constant for the two sexes. In that event, it is possible to write

$$\mu_{m}\,(a + x) = K(a, b)\,\mu_{f}\,(b + x) \tag{13}$$

where

$$K(a, b) = (B_m/B_f)\, C^{a-b} \tag{14}$$

in which $B_m$ and $B_f$ are the values of the parameter $B$ in (11) for the males and the females respectively. Substituting (13) in (10) and simplifying, the equation

$$P(a, b) = \frac{1}{1 + K(a, b)} = \frac{1}{1 + (B_m/B_f)\, C^{a-b}} \tag{15}$$

is obtained. It is of considerable interest to note from (15) that, in a given population, the probability of losing a spouse depends primarily on the age difference of the two spouses, but not on their specific ages. From a mathematical point of view, this surprising finding follows from the reasonable assumption of the Gompertz law of mortality with an additional (but reasonable nonetheless) restriction that for any age the forces of mortality for the two sexes remain proportional to one another. It is easy to see that minor deviations from these assumptions will not drastically affect the aforementioned results.

Also of interest to note is that the conclusions drawn about the behavior of the eventual probability measures apply also to the same calculated over a shorter time interval. In other words, the probability, say, $P_t\,(a, b)$, that the wife will be the first to die within the next $t$ years, conditional to at least one of them dying in that interval, will be the same as shown in (15). This is so because changing the upper limits of the integrals in (10) from $\alpha\,(a, b)$ to some $t$, has no effect on its value when Gompertz law of mortality is assumed. The unconditional probabilities (Depoid, 1938), as mentioned earlier, will increase with $t$ and approach the eventual probability as the limiting value.

## Empirical Estimates of $K(a, b)$

As shown in (15) $K(a, b)$ consists of two factors, namely, $B_m/B_f$ and $C^{a-b}$, in which the former can be expressed as the ratio of the forces of mortality of the corresponding sexes when both spouses are of the same age. That is to say,

$$\frac{B_m}{B_f} = \frac{\mu_m(x)}{\mu_f(x)}. \tag{16}$$

Since, in practice, the ratio will show some variation by age, an estimate of the same can be obtained as

$$\frac{B_m}{B_f} = \frac{\int \mu_m(x)dx}{\int \mu_f(x)dx} \tag{17}$$

where the limits of the integrals may be set at convenience. From practical considerations, the lower limit of the integral may be set at age 30 whereas the upper limit may be determined by the lower boundary, say $\alpha$, of the last interval (open ended) of the life tables. In that case (17) can be simplified as

$$\frac{B_m}{B_f} = \frac{ln\,[l_m\,(30)/l_m\,(\alpha)]}{ln\,[l_f\,(30)/l_f(\alpha)]}\,. \tag{18}$$

Similarly, the parameter $C$ which can be expressed either as

$$C^n = \frac{\mu_m\,(x+n)}{\mu_m\,(x)} \tag{19}$$

or as

$$C^n = \frac{\mu_f\,(x+n)}{\mu_f\,(x)} \tag{20}$$

will also show some variation by age and sex. Accordingly, an estimate of $C$ may be generated from

$$C^n = \frac{1}{2}\left[\frac{ln\,[l_m\,(30+n)/l_m\,(\alpha)]}{ln\,[l_m\,(30)/l_m\,(\alpha-n)]} + \frac{ln\,[l_f\,(30+n)/l_f\,(\alpha)]}{ln\,[l_f\,(30)/l_f\,(\alpha-n)]}\right]. \tag{21}$$

The values of $K\,(a,\,b)$ for different combinations of $a$ and $b$ can then be obtained from (14) through appropriate substitutions. These, for the 1973 U.S. Life Tables ($n = 5$ and $\alpha = 85$), are shown in Table 2.

TABLE 2—ESTIMATED VALUES OF $K(a,\,b) = (B_m/B_f)\,C^{a-b}$ FOR DIFFERENT VALUES OF HUSBAND WIFE AGE DIFFERENTIALS $a - b$

| $\mid a-b \mid$ | *0* | *1* | *2* | *3* | *4* | *5* | *6* | *7* | *8* | *9* | *10* |
|---|---|---|---|---|---|---|---|---|---|---|---|
| $K(a,b)$ $a \geqslant b$ | 1.71 | 1.86 | 2.02 | 2.20 | 2.40 | 2.61 | 2.84 | 3.09 | 3.36 | 3.66 | 3.99 |
| $K(a,b)$ $a \leqslant b$ | 1.71 | 1.57 | 1.44 | 1.32 | 1.22 | 1.12 | 1.02 | 0.94 | 0.87 | 0.80 | 0.73 |

SOURCE: 1973 U.S. Life Tables.

### Solution of $P(a,\,b)$ by Numerical Method

The integrals appearing in (10) can of course be evaluated by numerical methods which, as such, will be free from assumptions about the forces of mortality made earlier. First, it may be noted that from the definition of the force of mortality given in (4), it is possible to write

$$l_f\,(b + x)\,\mu_f\,(b + x) = -dl_f\,(b + x) \tag{22}$$

so that the numerator of (10) can be rewritten as

$$-\int_0^{\alpha\,(a,\,b)} l_m\,(a + x)\,dl_f\,(b + x). \tag{23}$$

In general, the derivative of $l(x)$ can be assumed as constant over a small age interval and therefore, for such an interval of length $n$ (usually no greater than 5) years

$$-\frac{d}{dx}\,l_f\,(b + x) = \frac{l_f\,(b + x) - l_f\,(b + x + n)}{n} = \frac{{}_nd_f\,(b + x)}{n} \tag{24}$$

is the average annual female deaths in the age interval $b + x$ to $b + x + n$. Therefore, (23) simplifies into

$$H_m\,(a, b) = \frac{1}{n}\sum_{i=0}\,{}_nL_m\,(a + in)\,{}_nd_f(b + x) \tag{25}$$

where

$${}_nL_m\,(a + in) = \int_{in}^{(i+1)n} l_m(a + x)\,dx \tag{26}$$

is the size of the stationary population in the age interval $a + in$ to $a + (i + 1)n$. Unfortunately, the upper limit of $i$ needed for the evaluation of (25) is not known as the life table functions are generally not available beyond some age $\alpha$ (80 or 85). Consequently, assumptions have to be made about the contributions of the terms in (25) beyond $\alpha$. It appears though, that when $a$ and $b$ are both sufficiently smaller than $\alpha$, $P\,(a, b)$, which because of (25) and a similar definition of $H_f\,(a, b)$ reduces to

$$P\,(a, b) = \frac{H_m\,(a, b)}{H_m\,(a, b) + H_f\,(a, b)} \tag{27}$$

after appropriate substitutions in (9), can be assumed to remain unaffected when the $H$ functions are evaluated over the age interval $(a, \alpha)$ for $a \geqslant b$ or over $(b, \alpha)$ when $a < b$. Although, it is not immediately apparent, such an assumption leads to the mathematical equality

$$P(a, b) = P(\alpha, \alpha - a + b) \tag{28}$$

for $a \geqslant b$, and

$$P(a, b) = P(\alpha - a + b, \alpha) \tag{29}$$

otherwise. These equalities can be established by first distinguishing the $H$ functions over the reduced interval as $H^{\alpha}$. In that case, for $a > b$, (27) can be written as

$$P(a, b) = \frac{H_m^{\alpha}(a, b) + H_m(\alpha, \alpha - a + b)}{H_m^{\alpha}(a, b) + H_f^{\alpha}(a, b) + H_m(\alpha, \alpha - a + b) + H_f(\alpha, \alpha - a + b)}. \tag{30}$$

Thus, the assumption of the equality

$$P(a, b) = \frac{H_m^{\alpha}(a, b)}{H_m^{\alpha}(a, b) + H_f^{\alpha}(a, b)} \tag{31}$$

produces the other equality, namely,

$$P(a, b) = \frac{H_m(\alpha, \alpha - a + b)}{H_m(\alpha - a + b) + H_f(\alpha - a + b)} \tag{32}$$

which also equals $P(\alpha, \alpha - a + b)$ because of (27). Similarly, (29) can be established for $a < b$. From logical considerations, it may be added that (28) and (29) further imply (but do not mathematically require) that

$$P(a, b) = P(a + h, b + h) \tag{33}$$

for all $h$. This is so, because like $P(a, b)$, the probability $P(a + h, b + h)$ can also be estimated from $H^{\alpha}(a + h, b + h)$ at least for small $h$, in which case

$$P(a + h, b + h) = P(\alpha, \alpha - a + b) \quad \text{for } a > b. \tag{34}$$

Because of (32), therefore, the equality proposed in (33) should also hold.

It is obvious that if $P(a, b)$ is estimated from (32) for a given age combination $a$ and $b$ of the spouses and the same is set equal to $P(\alpha, \alpha - a + b)$, the substitution of the latter in (6), where $a$ and $b$ are changed respectively to $a + h$ and $b + h$, will generate the estimates of $P(a + h, b + h)$ as

$$P(a + h, b + h) = \frac{H_m^\alpha(a + h, b + h) + l_m(\alpha)\, l_f(\alpha - a + b)\, P(\alpha, \alpha - a + b)}{l_m(a + h)\, l_f(b + h)}. \quad (35)$$

Needless to say, the mathematical equality of $P(a, b)$ and $P(a + h, b + h)$ does not follow from such a procedure, however, as mentioned earlier, the difference between these two probabilities should be negligible.

The reader must have noted the equivalence of the end results generated by the present method with those based on the Gompertz law of mortality. Surprising as the results may be, the eventual probability of becoming a widow or a widower (and similarly the conditional probability over any time interval), seems to be determined by the age difference of the spouses and not by their actual ages.

**Application on U.S. Data and Discussion of Results**

The values of $P(a, b)$ have been calculated from the 1973 U.S. Life Tables, on the basis of the two methods presented in this paper. These are shown in Table 3, in which the age difference between the two spouses has an arbitrarily chosen range of —10 to 10 years. It may be recalled that the principal parameters of the Gompertz model were estimated from the age interval (30, 85), from which $K(a, b)$ values were obtained and shown in Table 2. Substitutions of these values in (15) provide estimates of $P(a, b)$ which are shown in cols (2-3) of Table 3. Next, the $H^\alpha(a, b)$ functions, required for the method based on numerical integration, are obtained for those combinations of $a$ and $b$, such that the

$$\text{minimum } (a, b) = 15. \quad (36)$$

The choice of age 15 is based on a reasonable minimum of the observed ages of marriage, and in this way the difference between $\alpha$ and the minimum $(a, b)$ is maximized in order to strengthen the assumption resulting in (31) and (32). From the same life tables, these $H$ functions are then obtained for $\alpha = 85$, for substitutions in (31) to generate the estimates of $P(a, b)$ for different age combinations of the two spouses.

As expected, the values of $P(a + h, b + h)$ are found to be virtually invariant with respect to $h$, and instead of reproducing all such values, only the minimum

TABLE 3—PROBABILITY $P(a, b)$ OF HUSBAND $a$ YEARS OLD, OUTLIVING THE WIFE AGED $b$ YEARS, ESTIMATED BY THE METHODS BASED ON (1) GOMPERTZ LAW OF MORTALITY AND (2) NUMERICAL INTEGRATION

| $\|a-b\|$ | *P(a, b) by Gompertz Law* | | *Optimum values of P(a, b) by numerical integration* | | | |
|---|---|---|---|---|---|---|
| | | | $a \geqslant b$ | | $a \leqslant b$ | |
| | $a > b$ | $a < b$ | *Minimum* | *Maximum* | *Minimum* | *Maximum* |
| *(1)* | *(2)* | *(3)* | *(4)* | *(5)* | *(6)* | *(7)* |
| 0 | .37 | .37 | .35 | .38 | .35 | .38 |
| 1 | .35 | .39 | .33 | .36 | .36 | .40 |
| 2 | .33 | .41 | .31 | .34 | .38 | .41 |
| 3 | .31 | .43 | .29 | .32 | .40 | .43 |
| 4 | .30 | .45 | .27 | .30 | .42 | .45 |
| 5 | .28 | .47 | .26 | .28 | .44 | .47 |
| 6 | .26 | .49 | .24 | .26 | .46 | .49 |
| 7 | .25 | .51 | .22 | .24 | .48 | .51 |
| 8 | .23 | .53 | .21 | .23 | .50 | .53 |
| 9 | .22 | .55 | .20 | .21 | .52 | .55 |
| 10 | .21 | .57 | .18 | .20 | .55 | .57 |

SOURCE : 1973 U. S. Life Tables.

and the maximum values have been shown in cols (4-7) of Table 3 for integral values of $| a - b | \leqslant 10$.

The closeness of the estimates generated by the two different methods are mutually reinforcing for the validity of the separate assumptions on which these are based. According to the tabled values, the probability of becoming a widow is at least twice as large as that of becoming a widower when the husband is two to three years older than the wife. The differential risks of losing a spouse for this currently normative age difference is worth noting. The two probabilities become equal when the husband is about seven years younger than the wife, a figure which is slightly less than the difference between the life expectancies of the two sexes. It will be interesting to see how these probability measures compare with those generated from other life tables.

## Acknowledgement

I am grateful to Fr. Joseph C. O'Hara and Mr. S. N. Banerjee, both of the Department of Sociology and Anthropology for developing the computer program and for preparing the tables presented in the paper.

## References

1. Depoid, Pierre, 1938, Tables nouvelles relatives a la population francaise. *Bulletin de la statistique generals de la France,* 27(11), 269-324.
2. **Jordan, C. W-, Jr.,** 1967, *Life* ***Contingencies.*** **The** Society of Actuaries, **Chicago.**
3. **Spiegelman, M., 1969,** ***Introduction to Demography,*** Cambridge, Mass.
4. **USDHEW, 1976, Vital** Statistics **of the** United States, 1973, **Vol. II, Section 5, Life Tables.**