ABSTRACT

On the basis of experimental observations on animals, applications to clinical data on patients and theoretical statistical reasoning, the author developed a computer-assisted general mathematical model of the “probacent”-probability equation, Eq.1 and death rate (mortality probability) equation, Eq.2 derivable from Eq.1 that may be applicable as a general approximation method to make useful predictions of probable outcomes in a variety of biomedical phenomena [1-4]. Eqs.1 and 2 contain a constant, γ and c, respectively. In the previous studies, the author used the least maximum-difference principle to determine these constants that were expected to best fit reported data, minimizing the deviation. In this study, the author uses the method of computer-assisted least sum of squares to determine the constants, γ and c in constructing the “probacent”-related formulas best fitting the NCHS-reported data on survival probabilities and death rates in the US total adult population for 2001. The results of this study reveal that the method of computer-assisted mathematical analysis with the least sum of squares seems to be simple, more accurate, convenient and preferable than the previously used least maximum-difference principle, and better fitting the NCHS-reported data on survival probabilities and death rates in the US total adult population. The computer program of curved regression for the “probacent”-probability and death rate equations, may be helpful in research in biomedicine.

Keywords: Linear Regression; Curved Regression; Least Sum of Squares; Least Maximum-Difference; “probacent”-Probability Equation; Computer Program of Curved Regression; Survival Probability Equation; Death Rate equation; Mortality Probability; Human Tolerance to Radiation

1. INTRODUCTION

On the basis of experimental observations on animals, clinical applications on patients and theoretical statistical reasoning, the author developed a general mathematical model of “probacent”-probability equation that may be applicable as a general approximation method to make useful predictions of probable outcomes in a variety of biomedical phenomena [1-4].

The model of the “probacent”-probability equation was constructed from experimental studies on animals to express survival probability in mice exposed to g-force in terms of magnitude of acceleration and exposure time [1,5]; and to express a relationship among intensity of stimulus or environmental agent (such as drug [1,2,6], heat [7], pH [8], electroshock [7,9] and radiation [4,10]), duration of exposure and biological response in animals.

The model has been applied to data in the literature to express carboxyhemoglobin levels of blood as a function of carbon monoxide concentration in air and duration of exposure [11,12]; to express a relationship among plasma acetaminophen concentration, time after ingestion and occurrence of hepatotoxicity in man [13,14]; to predict survival probability in patients with malignant melanoma [15-17]; to express survival probability in patients with heart transplantation [18,19]; to express a relationship among age, height and weight, and percentile in Saudi and US children of 6 - 16 years of age [20-22]; to predict the percentile of heart weight by body weight from birth to 19 years of age [23,24]; and to predict the percentile of serum cholesterol levels by age in adults [25-27].

The model was applied to the United States life tables, 1992 and 2001 reported by the National Center for Health Statistics (NCHS) to construct formulas expressing age-specific survival probability, death rate and life expectancy in US adults, men and women [3,28-31].

The formula of survival probability is expressed by the following “probacent”-probability Eq.1:

\[ P^* = A - B \log T \]  

(1a)
where \( T \) = time after biomedical insult, diagnosis of cancer or age; \( P = \) “probacent” (abbreviation of probability percentage) = relative biological amount of “reserve” for survival; “probacent” \( (P) \) of \( 0, 50 \) and \( 100 \) corresponds to \(-5\) SD, mean and mean +5SD, respectively; the unit of “probacent” is 0.1 SD. In addition, 0, 50 and 100 “probacents” seem to correspond to 0, 50 and 100 percent probability in mathematical prediction problems in terms of percentage. Therefore, it seems to the author that survival probabilities can be used to predict probabilities in general biomedical phenomena. “probacent” \( (P) \) values are obtainable from a list of conversion of percent probability into “probacent” that was published by the author (Table 6 of Ref. [1] and Table of 4 of Ref. [2]). \( \gamma, A \) and \( B \) are constants; \( A \) is an intercept and \( B \) a slope; \( \gamma \) represents a curvature (a shape of curve) and expressed by the following equation:

\[
\gamma = \log \left( \frac{A - B \log T}{\log P} \right)
\]

If the value of \( \gamma \) becomes equal to one, Eq.1 represents a log-normal distribution. Eq.1 is considered to be fundamentally based on the Gaussian normal distribution.

Eq.2 representing death rate is derived from Eq.1 expressing survival probability [30].

\[
(\log D)^c = a + b \log T
\]

where \( D \) represents death rate in percentage (mortality probability); \( T \) is time or age; \( c, a \) and \( b \) are constants; \( c \) represents a curvature (a shape of curve) like \( \gamma \) in Eq.1a; \( a \) is an intercept and \( b \) a slope.

If the value of constant \( c \) becomes equal to one, Eq.2 is essentially similar to the Weibull distribution [32].

Eq.2 was applied to express death rates in US adults [3,30,31]. It was found to better express death rates in US total elderly population than the Gompertz, the exponential and the Weibull distributions [3].

Eq.2 has been successfully applied to predict mortality probability in total body irradiation without medical support in humans as a function of dose rate of radiation and duration of exposure [4], and to express mean survival time as a function of daily dose rate of total body irradiation in mice [33].

Mehta and Joshi [34] successfully applied the “probacent”-probability equation, Eqs.1 and 2 to use model-derived data as an input for radiation risk evaluation of Indian adult population.

The Constants, \( \gamma \) in Eq.1 of Survival Probability, and \( c \) in Eq.2 of Death Rate

If the constants, \( \gamma \) in Eqs.1 and \( c \) in Eq.2 are one, then both equations represent a straight line when data points are plotted against age on a graph paper as illustrated in Figures 1 and 2. If the \( \gamma \) and \( c \) values are >1, it indicates that the data-points-connecting curve would reveal an upward convexity by graphical inspection. If the \( \gamma \) and \( c \) values are <1, it indicates that the data curve would reveal a downward convexity on the graph.

The author used a principle of least maximum-difference, \( I(E-O) \) in determining the best-fitting \( \gamma \) and \( c \) values to the observed data curve. Here \( E \) and \( O \) in the parenthesis stand for formula-derived \( \gamma \) and \( c \) values.
Figure 2. Relationship between age and death rate in the US total elderly population of 60 - 100 years for 2001. The abscissa represents age in years and the ordinate death rate (D) in percentages (log scale). Data points of closed circles indicate US national life table death rates reported by the National Center for Health Statistics (NCHS) for 2001. The dashed straight line represents death rates predicted by the Gompertz mortality model expressed by equation,

\[ D = 10 \left( -2.2674 + 0.03779 T \right) \]

The solid curved line represents death rates predicted by the “probacent”-probability model of death rate \( D \) expressed by Eqs.7 and 8. Data points of NCHS appear to fall overall on the solid death-rate line predicted by Eqs.7 and 8. The maximum predictive error of the “probacent” model is ±0.3% and that of the Gompertz model ±3.2%. Source: reference [3].

In analysis of the least maximum-difference, random different values of integer and/or fractional number are substituted as \( \gamma \) and \( c \) values in Eq.1 or 2 to calculate survival probabilities, \( S \) or death rates, \( D \). The above described method of the least maximum-difference principle was used in the author’s previous publications to minimize the deviation. The least sum of squares of well-known linear regression in statistics [32,35,36] is not employed in the previous author’s studies. However, to my knowledge, there seem to be no computer-program-assisted, nonlinear, curved regression models of the least sum of squares in the literature that determine the best-fitting constant, \( \gamma \) or \( c \) value in the “probacent”-probability or death rate equation, Eq.1 or 2, minimizing the sum of deviation [37-42].

The purpose of this study is to design a computer program of nonlinear, curved regression of the least sum of squares for construction of best-fitting equations. of “probacent”-probability and death rate developed by the author to the NCHS-reported data [29].

2. MATERIALS AND METHODS


The author published computer-assisted predictive formulas expressing the NCHS-reported survival probabilities, death rates (mortality probabilities) and life expectancies in US adults, men and women, 2001, employing a model of the “probacent”-probability and death-rate equations previously published by the author in the study [3]. The survival probability is percent probability of surviving to the beginning of age \( T \) from birth. The death rate is percent probability of dying between age \( T \) to \( T + 1 \).

The data are plotted on a log-log graph paper as illustrated in Figures 1 and 2.

In this study, the data on survival probabilities and death rates shown in the NCHS’ report [29] and [3] as well as Figures 1 and 2 are used to design computer programs of nonlinear, curved regression of the least sum of squares for the “probacent”-probability and death rate equations to minimize the sum of deviations, and to find the best-fitting constant values, \( \gamma \) and \( c \).

2.1. Use of the Least-Maximum-Difference Principle in Analysis

In the author’s previous studies, the least maximum-difference principle, least \( |E-O| \) (the absolute value of the difference) is used to minimize the deviation.

2.1.1. Formulas of Survival Probabilities (S)

A mathematical method to determine constants, \( \gamma \), \( A \) and \( B \) in Eq.1 is described in Appendix of Ref. [3].

Two sets of data on age (\( T \)) and survival probability (S) are used in each age group, 20 - 60, 60 - 85 or 85 - 100 years to determine constants \( A \) and \( B \) as seen in Eqs.3a, 3b and 3c, respectively. The most appropriate and best-fitting \( \gamma \) values of Eq.1 for the age groups of 20 - 60, 60 - 85, and 85 - 100 years are determined, using the least maximum-difference principle and comparing maximum
differences I (E-O) I calculated by substituting a various semi-random and semi-selective values as the γ value in Eqs.3a, 3b and 3c.

\[ P^\gamma = 4.67677 \times 71.002^\gamma - 3.67677 \times 61.605^\gamma \]
\[ -2.63013 \times \left( 71.002^\gamma - 61.605^\gamma \right) \times \log T \]  
\[ (3a) \]

\[ P^\gamma = 12.75482 \times 61.605^\gamma - 11.75482 \times 46.405^\gamma \]
\[ -6.6107 \times \left( 61.605^\gamma - 46.405^\gamma \right) \times \log T \]  
\[ (3b) \]

\[ P^\gamma = 28.33664 \times 46.405^\gamma - 27.33664 \times 29.538^\gamma \]
\[ -14.16832 \times \left( 46.405^\gamma - 29.538^\gamma \right) \times \log T \]  
\[ (3c) \]

The following Eqs. 4, 5 and 6 are thus constructed to express survival probabilities of the three age groups:

The age group of 20 - 60 years: Eqs.4a and 4b.

\[ P^{2,7} = 4.67677 \times 71.002^{2,7} - 3.67677 \times 61.605^{2,7} \]
\[ -2.63013 \times \left( 71.002^{2,7} - 61.605^{2,7} \right) \times \log T \]  
\[ (4a) \]

The age group of 60-85 years: Eqs.5a and 5b.

\[ P^{4,8} = 12.75482 \times 61.605^{4,8} - 11.75482 \times 46.405^{4,8} \]
\[ -6.6107 \times \left( 61.605^{4,8} - 46.405^{4,8} \right) \times \log T \]  
\[ (5a) \]

The age group of 85-100 years: Eqs.6a and 6b.

\[ P^{2,3} = 28.33664 \times 46.405^{2,3} - 27.33664 \times 29.538^{2,3} \]
\[ -14.16832 \times \left( 46.405^{2,3} - 29.538^{2,3} \right) \times \log T \]  
\[ (6a) \]

\[ S = \frac{10}{\sqrt{2\pi}} \int_0^\infty \exp \left[ \frac{-(P-50)^2}{200} \right] dP \]  
\[ (4b, 5b, 6b) \]

2.1.2. Formulas of Death Rates (D)

Constants, c, a and b are determined likewise as above described (the author’s note: see Appendix of Reference [33] if needed) and the following equations are constructed to express death rates of the two age groups:

The age group of 60 - 85 years:

\[ \log D^{0.82} = 12.75481 \times 0.00655^{0.82} - 11.75481 \times 0.97102^{0.82} \]
\[ + 6.6107 \times \left( 0.97102^{0.82} - 0.00655^{0.82} \right) \times \log T \]  
\[ (7) \]

The age group of 85 - 100 years:

\[ \log D^{1.7} = 30.13651 \times 0.97102^{1.7} - 29.13651 \times 1.42545^{1.7} \]
\[ + 15.10118 \times \left( 1.42545^{1.7} - 0.97102^{1.7} \right) \times \log T \]  
\[ (8) \]

2.2. Use of the Least Sum of Squares in Analysis

In this study, the least sum of squares is used.

2.2.1. Formulas of Survival Probabilities (S)

The method of least sum of squares, least \( \sum (E-O)^2 \) is used to determine the best-fitting \( \gamma \) and \( c \) values of the “probacent”-probability equation to minimize the sum of deviations. Abridged five-year intervals are used for analysis to simplify computer programs.

A close look at the data points in Figure 1 in graphic inspection suggests that the line connecting data points at each age group of 20 - 60, 60 - 85 and 85 - 100 years bulges upward, revealing an upward convexity and so that the \( \gamma \) value is >1. If the line shows a straight line, it indicates \( \gamma = 1 \). If the line reveals a downward like the line connecting the data points on death rates of the age group of 60 - 85 years in Figure 2, it would indicate 0 < \( \gamma < 1 \).

A three-step approach in analyzing data with help of the computer program is taken to find the best-fitting constant values, \( \gamma \) and \( c \) in Eqs.1 and 2.

The first step of computer-assisted mathematical analysis:

Enter an integer \( N \), starting from 1 and increasing the integer, 2, 3, … up to \( N \) as the \( \gamma \) value in Eq.3a for the age group of 20 - 60 years in US adults. Sums of squares, \( \Sigma (E-O)^2 \) are calculated with the computer program shown in Figure 3. The computer-derived line representing Eq.3 with a specific \( \gamma \) value of 1 to \( N \) first approaches toward the NCHS-reported-data line from the starting straight line; the sum of squares would be gradually decreasing. When the computer-generated line touches the NCHS-reported-data line, the sum of squares becomes minimum, the least sum, ideally zero. After passing the NCHS-data line, the sum of squares with increasing \( \gamma \) values would suddenly begin to increase and continues to increase further more. These processes are shown in Table 1.

The second step of computer-assisted mathematical analysis:

If the sum of squares suddenly starts increasing after preceding gradual decrease at integer \( N + 1 \) of \( \gamma \) value, then enter \( N - 0.1 \) and \( N + 0.1 \) as \( \gamma \) value in Eq.3a. Calculate the sums of squares. Compare the sums at \( N - 0.1 \) and \( N + 0.1 \) with the sum at \( N \).

The third step of computer-assisted mathematical analysis:

If the sum at \( N - 0.1 \) is smaller than the sum at \( N \), then enter \( (N - 1) + 0.1 \), \( (N - 1) + 0.2 \), … \( (N - 1) + 0.9 \) as \( \gamma \) value in Eq.3a. Compare the sums of squares and choose the number with the least sum of squares that is determined to be the best-fitting \( \gamma \) value for Eq.3a. A very close and best agreement is found between the computer-derived and NCHS-reported survival probabilities with the \( \gamma \) value of 12.8. Eqs.9a and 9b, are finally derived to best represent a relationship between
Figure 3. The computer program to calculate the sum of squares, $\Sigma (E-O)^2$ as a function of $\gamma$ value and age ($T$) in the US total adult population. Results of execution of the program are shown in Tables 1 and 3. This program is for $\gamma$ value of 12.8 in Eq.4a for the age group of 20 - 60 years.
The equations of survival probabilities, Eqs. 10a and 11, for the age groups of 60 - 85 and 85 - 100 years are likewise derived as shown in Table 1.

The age group of 20 - 60 years: Eqs. 9a and 9b

\[
P^{12.8} = 4.67677 \times 71.002^{12.8} - 3.67677 \times 61.605^{12.8} - 2.63013 \times (71.002^{12.8} - 61.605^{12.8}) \times \log T \quad (9a)
\]

The age group of 60-85 years: Eqs. 10a and 10b.

\[
P^{4.8} = 12.75482 \times 61.605^{4.8} - 11.75482 \times 46.405^{4.8} - 6.6107 \times (61.605^{4.8} - 46.405^{4.8}) \times \log T \quad (10a)
\]

The age group of 85-100 years: Eqs. 11a and 11b.

\[
P^{2.3} = 28.33664 \times 46.405^{2.3} - 27.33664 \times 29.538^{2.3} - 14.16832 \times (46.405^{2.3} - 29.538^{2.3}) \times \log T \quad (11a)
\]

\[
S = \frac{10}{\sqrt{2\pi}} \int_{-\infty}^{P-50} \exp \left[-\frac{(P-50)^2}{200}\right] dP \quad (9b, 10b, 11b)
\]

Both methods of mathematical analysis, the least-maximum-difference and the least sum of squares give different $\gamma$ values, 1.7 and 1.8 for the age groups of 20 - 60 years. However, both methods give same $\gamma$ values, 4.8 and 4.8 for the age group of 60 - 85 years, and 2.3 and 2.3 for the Age Group of 85 - 100 Years.

### 2.2.2. Formulas of Death Rates (D)

The constants $c$, $a$ and $b$ are likewise derived as explained above and as seen in Table 2. Fractional numbers are used to determine these constants. Two following formulas expressing death rates for the age groups of 60 - 85 and 85 - 100 years for the US total elderly population:

The age group of 60-85 years: Eq. 12.

\[
\log(D)^{0.79} = 12.75481 \times 0.00655^{0.79} - 11.75481 \times 0.97102^{0.79} + 6.6107 \times (0.97102^{0.79} - 0.00655^{0.79}) \times \log T
\]

The age group of 85 - 100 years, Eq. 13.

\[
\log(D)^{1.8} = 30.13651 \times 0.97102^{1.8} - 29.13651 \times 1.42545^{1.8} + 15.10118 \times (1.42545^{1.8} - 0.97102^{1.8}) \times \log T
\]

---

**Table 1.** Sums of squares of differences, $\Sigma (E-O)^2$ in nonlinear, curved regression of the least sum of squares to determine a best-fitting $\gamma$ value for the “probacent”-probability equation expressing age-specific survival probabilities $(S)^a$ in US total adult population. Sums of squares of differences are calculated by computer programs. A representative program is illustrated in Figure 3.

<table>
<thead>
<tr>
<th>Age group</th>
<th>20 - 60 years</th>
<th>60 - 85 years</th>
<th>85 - 100 years</th>
</tr>
</thead>
<tbody>
<tr>
<td>Used “probacent” equation</td>
<td>Eq.3a</td>
<td>Eq.3b</td>
<td>Eq.3c</td>
</tr>
<tr>
<td>Finally chosen $\gamma$ value (N)</td>
<td>12.8</td>
<td>4.8</td>
<td>2.3</td>
</tr>
<tr>
<td>$N$</td>
<td>$\Sigma (E-O)^2$ change</td>
<td>$N$</td>
<td>$\Sigma (E-O)^2$ change</td>
</tr>
<tr>
<td>1</td>
<td>19.211494</td>
<td>1</td>
<td>141.460900</td>
</tr>
<tr>
<td>2</td>
<td>16.037229</td>
<td>D**</td>
<td>2</td>
</tr>
<tr>
<td>3</td>
<td>13.142712</td>
<td>D</td>
<td>3</td>
</tr>
<tr>
<td>4 – 11 continue to decrease</td>
<td>4</td>
<td>6.421165</td>
<td>D</td>
</tr>
<tr>
<td>12</td>
<td>0.148060</td>
<td>D</td>
<td>5</td>
</tr>
<tr>
<td>13</td>
<td>0.086081</td>
<td>D</td>
<td>6</td>
</tr>
<tr>
<td>14</td>
<td>0.263210</td>
<td>I***</td>
<td>4.9</td>
</tr>
<tr>
<td>12.9</td>
<td>0.081272</td>
<td>#</td>
<td>5.1</td>
</tr>
<tr>
<td>13.1</td>
<td>0.093282</td>
<td>D</td>
<td>4.1</td>
</tr>
<tr>
<td>12.1</td>
<td>0.130686</td>
<td>D</td>
<td>4.2</td>
</tr>
<tr>
<td>12.2</td>
<td>0.115832</td>
<td>D</td>
<td>4.3</td>
</tr>
<tr>
<td>12.3</td>
<td>0.103485</td>
<td>D</td>
<td>4.4</td>
</tr>
<tr>
<td>12.4</td>
<td>0.093632</td>
<td>D</td>
<td>4.5</td>
</tr>
<tr>
<td>12.5</td>
<td>0.086258</td>
<td>D</td>
<td>4.6</td>
</tr>
<tr>
<td>12.6</td>
<td>0.081349</td>
<td>D</td>
<td>4.7</td>
</tr>
<tr>
<td>12.7</td>
<td>0.078891</td>
<td>D</td>
<td>(4.8)</td>
</tr>
<tr>
<td>12.8</td>
<td>(0.078871)</td>
<td>D</td>
<td>4.9</td>
</tr>
<tr>
<td>12.9</td>
<td>0.081272</td>
<td>I</td>
<td>5</td>
</tr>
<tr>
<td>13</td>
<td>0.086081</td>
<td>I</td>
<td></td>
</tr>
</tbody>
</table>

---

$S = \frac{10}{\sqrt{2\pi}} \int_{-\infty}^{P-50} \exp \left[-\frac{(P-50)^2}{200}\right] dP$
Both methods of mathematical analysis, the least maximum-difference and the least sum of squares give different \( c \) values, 0.82 and 0.79 for the age group of 60-85 years, and 1.7 and 1.8 for the age group of 85-100 years, respectively.

### 2.3. Description of the Computer Program

The programs were written in UBASIC for IBM PC microcomputer and compatibles for Eqs.3-13. The computer program uses a formula of approximation instead of the integral of Eq.1b and Eqs.4b, 5b, 6b, 9b, 10b, 11b) because the computer cannot perform integral [2, 43-45]. Mathematical transformation of integral, Eq.1b to the formula of approximation is described in detail in the author’s book [45]. A representative computer program is illustrated in Figure 3 to calculate the sum of squares, \( \Sigma (E-O)^2 \) with the \( \gamma \) value of 12.8 in Eq.9a.

### 2.4. Statistical Analysis

A \( \chi^2 \) goodness-of-fit test (logrank test) [35] is used to test the fit of mathematical models to the NCHS-reported data [29]. The differences are considered statistically significant when \( p < 0.05 \).

### 3. RESULTS

Tables 3 and 4 show comparison of least maximum-differences, I(E-O)I, least sum of squares, \( \Sigma (E-O)^2 \) and \( \chi^2 \)-test \( p \) value in the two analytical methods of the least maximum-difference and least sum of squares, in age-specific survival probabilities and death rates for US total elderly population, calculated by computer programs as shown in a representative program, Figure 3.

The \( \gamma \) values in the survival probability equation in both methods are different, 12.7 and 12.8 in Eqs.4a and 9a for the age group of 20-60 years but same 4.8 and 4.8 in Eqs.5a and 10a for the age group of 60-85 years, 2.3 and 2.3 in Eqs.6a and 11a for the age group of 85-100 years. The \( c \) values in the death rate equation in both methods are all different, 0.82 and 0.79 in Eqs.5 and 10, and Eqs.6 and 11 for the age groups of 60-85 and 85-100 years, respectively.

The least maximum-difference and the least sum of squares reveal slightly smaller values in those in the least sum of squares than in the least maximum-difference but same values in Eqs.5 and 10, and Eqs.6 and 11 for the age groups of 60-85 and 85-100 years. The above results suggest that regression curves of the least sum of squares are closer to the NCHS-data-connecting line than those of the least maximum-difference. The \( \chi^2 \)-test \( p \) values are all >0.995, suggesting a very close agreement between both values of computer-derived and NCHS-reported survival probabilities and death rates.

The above described results seem to indicate that the analytical method of the least sum of squares is simpler, convenient and preferable, and give more accurate in
Table 3. Comparison of the least maximum-difference, $\|E-O\|$, the least sum of squares, $\sum (E-O)^2$ and $\chi^2$-test $p$ value in the two analytical methods of the least maximum-difference and the least sum of squares in age-specific survival probabilities for US total adult population.

<table>
<thead>
<tr>
<th>Age group</th>
<th>20 - 60 years</th>
<th>60 - 85 years</th>
<th>85 - 100 years</th>
</tr>
</thead>
<tbody>
<tr>
<td>Used “probacent” equation</td>
<td>Eq.4a</td>
<td>Eq.9a</td>
<td>Eq.5a</td>
</tr>
<tr>
<td>$\gamma$ value</td>
<td>12.7*</td>
<td>12.8**</td>
<td>4.8*</td>
</tr>
<tr>
<td>Least maximum-Difference, $|E-O|$</td>
<td>0.158</td>
<td>0.148</td>
<td>0.4</td>
</tr>
<tr>
<td>Least sum of Squares, $\sum (E-O)^2$</td>
<td>0.078891</td>
<td>0.078871</td>
<td>0.230143</td>
</tr>
<tr>
<td>$\chi^2$-test $p$ value</td>
<td>&gt;0.995</td>
<td>&gt;0.995</td>
<td>&gt;0.995</td>
</tr>
</tbody>
</table>

* $\gamma$ value is obtained by the method of the least maximum-difference, $\|E-O\|$. ** $\gamma$ value is obtained by the method of the least sum of squares of curved regression, $\sum (E-O)^2$. 'E' indicates computer-derived value of survival probability. 'O' indicates NCHS-reported value of survival probability [29] (see text).

Table 4. Comparison of the least maximum-difference, $\|E-O\|$, the least sum of squares, $\sum (E-O)^2$ and $\chi^2$-test $p$ value in the two analytical methods of the least maximum-difference and the least sum of squares in age-specific death rates for US total elderly population.

<table>
<thead>
<tr>
<th>Age group</th>
<th>60 - 85 years</th>
<th>85 - 100 years</th>
</tr>
</thead>
<tbody>
<tr>
<td>Used equation</td>
<td>Eq.10</td>
<td>Eq.12</td>
</tr>
<tr>
<td>$\gamma$ value</td>
<td>0.82*</td>
<td>0.79**</td>
</tr>
<tr>
<td>Least maximum-difference, $|E-O|$</td>
<td>0.359</td>
<td>0.325</td>
</tr>
<tr>
<td>Least sum of Squares, $\sum (E-O)^2$</td>
<td>0.064304</td>
<td>0.043004</td>
</tr>
<tr>
<td>$\chi^2$-test $p$ value</td>
<td>&gt;0.995</td>
<td>&gt;0.995</td>
</tr>
</tbody>
</table>

* $\gamma$ value is obtained by the method of the least maximum-difference principle, $\|E-O\|$. ** $\gamma$ value is obtained by the method of the least sum of squares of curved regression, $\sum (E-O)^2$. 'E' indicates computer-derived value of survival probability. 'O' indicates NCHS-reported value of survival probability [29] (see text).

determining values of $\gamma$ and $c$ constants in the “probacent”-probability and death rate equations.

4. DISCUSSION

Comparison of data shown in Tables 3 and 4 suggests a very close agreement between formula-derived and NCHS-reported data on survival probabilities and death rates in US total adult population because $\chi^2$-test $p$ values are >0.995 for each equation expressing them.

However, the method of the least sum of squares, least $\sum (E-O)^2$ gives more accurate and best fitting values of constants, $\gamma$ and $c$ in these equations that fit better the NCHS-reported data, closer to the data-points connecting line. The computer program of curved regression of the least sum of squares for the “probacent”-probability and death rate seems preferable to the method of the least maximum-difference, least $\|E-O\|$ to minimize the deviation.

The author feels that in a variety of biological phenomena, $\gamma$ and $c$ values are, if applicable, generally greater than one or less than one but not one, indicating a curved line when plotted on a X-Y graph paper as seen in Figures 1 and 2. The $\gamma$ and $c$ values are relatively rarely one, indicating a straight line on a graph or otherwise approximately appearing straight. This phenomena seems to be possibly analogous in physics to that light path is actually curved when passing through a gravitational field of space but appears straight [46,47].

If the $\gamma$ value becomes equal to one, Eq.1 represents a log-normal distribution. If the $c$ value is one, Eq.2 that is derivable from Eq.1 becomes essentially similar to the Weibull distribution [32]. Weibull distribution is a generalized exponential distribution [32]. If the base of a logarithm is one, the lognormal distribution would become a normal distribution (log $1^x = n$) [45, 48]. If the logarithm of one as its base is taken for X axis of time, the Gompertz distribution might be similar to the Weibull distribution. Therefore, it seems to the author that the Gompertz distribution might be a specific form of the “probacent”-probability equation. A normal distribution is likewise a specific form of the “probacent”-probability equation.

“probacent” can be a dependent variable versus an independent variable such as time or age as seen in survival probability, death rate and life expectancy in US total adult population (NCHS) [3,29]. “probacent” can be a dependent variable versus two independent variables such as intensity of stimulus or harmful agent and duration of exposure like dose rate of radiation and duration of exposure in total body irradiation [4], and like
dose of drug and time after administration [2,14]. In cases of two independent variables, Eq.1 can make a prediction of probability of occurrence of a response in subjects in various biomedical phenomena. The original and ultimate purpose of the author’s studies has been to find a general mathematical model, possibly a mathematical law hidden in nature that might calculate the probability of safe survival in humans and other living organisms exposed to any harmful or adverse circumstances, overcoming the risk [1,45].

The “probacent”-probability does not predict a single definite result or response for an individual observation in biodynamic biological phenomena. Instead, if the same observations are made on a large number of similar population, each of who had the same condition at the start, the model would predict the possible outcomes, the approximate biomedical events in quantities under observations, but it could not predict the occurrence of the specific event in an individual. Thus, the “probacent”-probability would introduce an unpredictability in biomedicine like an uncertainty principle of Werner Heisenberg in quantum mechanics [46,47].

The computer program represented by Figure 3 can easily calculate survival probabilities that are required to determine the least sum of squares, by using an approximation instead of integral in Eqs.4b, 5b, 6b, 9b, 10b, 11b. This enables users of the “probacent” model in mathematical analysis, to eliminate a need for consultation of table of normal frequency or percentile in books of statistics and mathematics.

5. CONCLUSIONS

In this study, a computer program of nonlinear, curved regression of the least sum of squares is designed to determine the constant values of γ in Eq.1 and c in Eq.2 that seems better fitting and more accurate than those obtained by the least maximum-difference principle as suggested by the data shown in Tables 3 and 4. The regression curve obtained by this method of the least sum of squares is closer to the data-point-connecting line than that obtained by the least maximum-difference principle. The computer program of curved regression for the “probacent”-probability equation may be helpful in research in biomedicine. The computer program of curved regression of this study would need further improvement to enable users to readily find the best-fitting constant values in the equations of the “probacent”-probability and death rate.

6. ACKNOWLEDGEMENTS

The author thanks Dr. C. W. Sheppard and Bruce Presley for their teachings in computer programming that made the author’s studies possibly published to the academic world.

REFERENCES

Clinic Proceedings I. (Growth): A quantitative anatomic study of 200 specimen human hearts during the first 10 decades of life. Part.


doi:10.1016/0020-7101(93)90015-X


doi:10.1016/0020-7101(94)90124-4


doi:10.1016/0020-7101(90)90049-Z


doi:10.1016/0020-7101(94)901068-C


doi:10.1016/S0169-2607(96)01778-6


