Austrian Journal of Statistics
2024, Volume 53, 64–84.
http://www.ajs.or.at/
doi:10.17713/ajs.v53i4.1833

About the Residual Time of Human Life

Kateryna Akbash
Volodymyr Vynnychenko Central Ukrainian State University
Kropyvnytskyi, Ukraine

Ivan Matsak
Taras Shevchenko National University of Kyiv
Kyiv, Ukraine

Abstract

Many researchers believe that human lifetime in the interval from 20 to 80 years is approximately described by the Gompertz distribution. It is natural to consider the problem of the distribution of the residual human lifetime at the final stage, that is, after 80 years. To do this, in this study, a statistical analysis of some statistical models on real data is carried out. Several models are presented, which, according to the authors, can adequately describe data on mortality after 80 years.

Keywords: residual, lifetime, statistical models.

1. Introduction

It is clear that various questions connected with the duration of life have always been of interest to mankind. It is known that such great mathematicians as K. Hugens, W. Leibniz, L. Euler, P. Laplace were interested in statistics about mortality and their presentation in the form of life tables.

So, back in 1760, L.Euler published a work called "General Researches on the Mortality and the Multiplication of the Human Race".

In 1812, P. Laplace wrote a classic course on probability theory, which laid the theoretical basis for life tables and gave a direct method for constructing them (see Gavrilov and Gavrilova (1991), Rosset (1981)).

How to quantify this phenomenon? And why is it necessary to study life expectancy? Here is the answer from biologists (Gavrilov and Gavrilova (1991)) and geneticists (Comfort (1964)). “...The practical significance of such studies is that they open up opportunities for predicting the life expectancy of organisms and, most importantly, the possibility of finding ways to continue life...”

“... Is it possible to understand what life is without finding out why it is limited in time and how its boundaries are determined? This is a fundamental problem of natural science, key to the entire scientific horizon”

Unfortunately, more than 250 years have passed since the time of Euler, when the first steps were taken in this problem, but so far there is much unknown here.

One of the first most successful attempts to mathematically describe the dependence of mortality on age belongs to the English actuaries B. Gompertz and W. Makeham. They, relying on statistical data and some simple empirical considerations, proposed to describe the human lifetime T using the following distribution function

$$\mathbf{P}(T < x) = F(x) = \begin{cases} 1 - \exp(-\lambda x - \frac{\beta}{\alpha}(\exp(\alpha x) - 1)) & \text{for} \quad x \geq 0 \\ 0, & \text{otherwise}, \end{cases} \tag{1}$$

where $x > 0$, $\beta > 0$, $\alpha > 0$.

The failure rate function (mortality intensities) for distribution (1) rows exponentially and has the following representation

$$r(x) = \frac{F'(x)}{1 - F(x)} = \lambda + \beta \exp(\alpha x), \quad x > 0. \tag{2}$$

B.Gompertz (Gompertz (1825)) considered formulas (1), (2) with $\lambda = 0$ (Gompertz (G) distribution). Subsequently, W.Makeham (Makeham (1860)) set the background component $\lambda > 0$ (Gompertz-Makeham (GM) distribution).

Formulas (1), (2) mean that mortality is influenced by two components: one that does not depend on age (background component) and the other, which at $\alpha > 0$ increases in geometric progression with age (characterizes exhaustion and aging of the organism).

Note that the Gompertz-Makeham law (1), (2) law turned out to be more competitive than many modern formulas. Somewhere in the interval after 20 years, formulas (1), (2) describe real data such as the lifespan of biological systems is better than other known formulas. The reason for this is simple, this law is closely related to the distribution of extreme values for minima (3rd type according to the classification of B.V.Gnedenko (Gnedenko (1943), Leadbetter, Lindgren, and Rootzen (1983))).

And what is known about the residual time of human life, for example, after 100 years?

Here is the opinion of biologists on this matter (Greenwood and Irwin (1939), Economos (1983), Gavrilov and Gavrilova (1991)):
"... An analysis of data on the mortality of people over 100 years old shows that in the most extreme age groups, the human mortality intensities practically ceases to increase with age, so that the kinetics of mortality of centenarians coincides with the kinetics of radioactive decay, and the half-life corresponds to approximately one year".

It is clear that at every age there is a certain probability of death. The data analysis carried out in work Thatcher (1999) hows that after 30 years this probability begins to increase by approximately 10% with each subsequent year. In round figures, the probability of death during the year (for modern men) increases from 0,001 at the age of 30 to 0,1 at the age of 80 years. This inevitable increase corresponds quite accurately to the "mortality law" of Gompertz. Somewhere after 80 years, the growth rate slows down, and there is still controversy about what happens at an older age.

It is well known that in the 19th century the probability of dying between the ages of 30 and 80 was much higher than today. Yet, until recently, it was widely believed that, despite impressive improvements at a younger age, the probability of dying over 80 years of age has not changed significantly. Thus, for example, article Modig, Andersson, Vaupel, Rau, and Ahlbom (2017) establishes that the mortality among centenarians (after 100 years) does not change, despite improvements at a younger age. The analysis was based on individual-level data for all Swedish and Danish centenarians who were born between 1870 and 1991; a total of 3006 men and 10963 women were included.

Some researchers believe that somewhere above 100 years there is a maximum life expectancy that has remained unchanged from ancient times (see, for example, Karin and de Haan

(1994) and the literature in Karin and de Haan (1994)). The opposite point of view and its justification can be found in Feller (1957), ch.1, § 1.

We also note monograph Thatcher, Kannisto, and Vaupel (1998), which considers six possible models of mortality for residual life after 80 years. Eight datasets were analyzed, each containing pooled data for 13 industrialized countries. The datasets cover periods 1960-1970, 1970-1980 and 1980-1990 years, as well as for cohorts born in 1871-1880 years, separately for men and women. They contain over 32 million deaths aged 80 and over during this period. The authors conclude that these data are clearly closer to the general logistic model and model Kannisto special case of the logistic model) than to the models of Gompertz, Weibull and Heligman & Pollard.

More precise mathematical statements on this topic were obtained in a recent article Rootzen and Zholud (2017). These authors, relying on real data from the database of centenarians IDL[1] about the residual life span of supercentenarians (the residual life of supercentenarians is their life time after 110 years) and on known statistical criteria, came to an important conclusion:
the residual time of human life after 110 years has an exponential distribution.

In this work, you can also find some other interesting information about the residual life span of supercentenarians, as well as a fairly complete bibliography in recent years on this topic. For example, in the work it is calculated that the probability of surviving another year after 110 years is 47%. The chances that a woman will live to 110 years are 2 in 100 000, and for men, these chances are 10 times less. Behind the data of base IDL, although there are many more female supercentenarians, it is shown that there are no differences between the distributions of women and men. There are also no significant differences in the distribution for the regions of Southern and Northern Europe, between North America and Japan and other countries. Based on the database data, the parameter of the exponential distribution of the residual lifespan of 1.34 and a confidence interval (1.22, 1.46) with a reliability of 0.95 were found.

Our calculations on the new additional data from IDL confirm the results of Holger Rootzen, Dmitrii Zholud. In addition, it follows from our analysis that the result about the exponential distribution for residual life ex‘pectancy remains true for men after 105 years. True hypothesis requires further testing.

On the other hand, for women in many cases this is not the case.

Here, just as in works Thatcher *et al.* (1998) and Rootzen and Zholud (2017), we pose the question:
what is the distribution of the residual duration of human life, for example, after 80 years?

Unfortunately, the above works do not provide an answer to this question with a reliability of 0.95 traditional for applied statistics.

From the plot of empirical mortality intensities for the US population (1980) (see Fig.1, empirical data was obtained from book Bowers, Gerber, Hickman, Jones, and Nesbitt (1997), Ch. 3, § 3.3, Table 3.3.1) that the exponential distribution cannot adequately describe the duration of human life after 80 years.

The situation with the Gompertz or Gompertz-Makeham distribution is more complicated and requires additional analysis.

In this work, we will try to answer these questions.

Our article has the following structure.

In subsection 2, one redundancy model from the mathematical reliability theory is studied, which, in our opinion, is directly related to our problem. From this model, one simple distribution (W^* - distribution) is derived, which will be used by us in the analysis of data such as residual life expectancy.

[1]International database on longevity (IDL).
URL: https://www.supercentenarians.org/en/data-and-metadata/

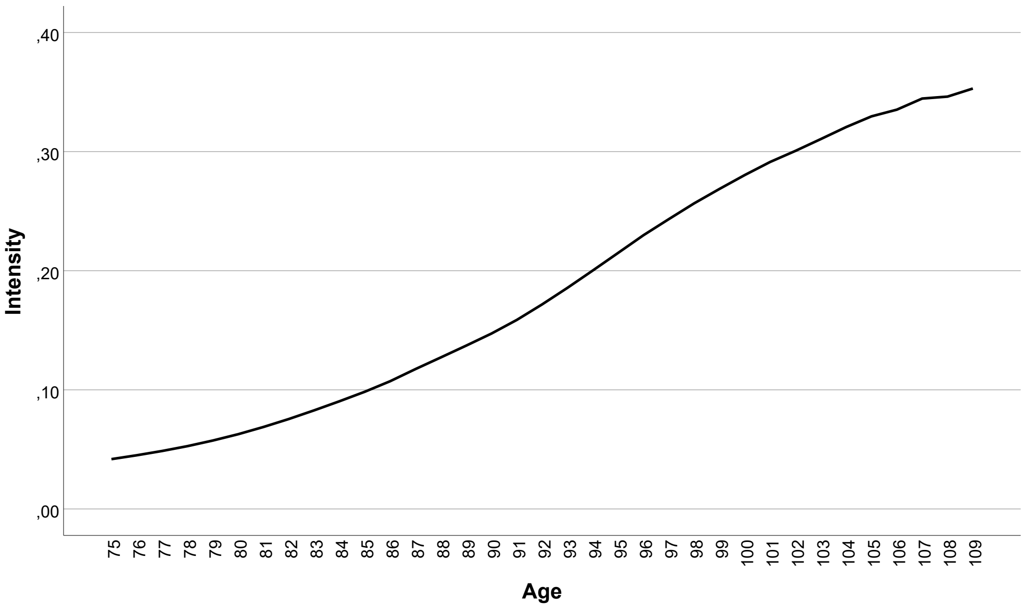

Figure 1: Mortality intensities after age 75 (USA)

And finally, in subsection 3 for the case of the W^* distribution, Gompertz (1) distribution and some others, a statistical analysis of real data for residual life expectancy after 80 years is carried out.

2. One mathematical model of reliability theory. Distribution W_m^*.

Mathematical models of reliability theory have long been used by biologists in the analysis of data on the mortality of people and other living organisms (see Gavrilov and Gavrilova (1991), ch. 6, § 5). At the same time, human life span is naturally considered at the main stage, that is, somewhere from 20 to 80 years. Also, an attempt is often made to derive the distribution laws of Gompertz or Weibull, or to explain their appearance more precisely, starting from some of the fundamental features of the biosystem.

We will consider a narrower task - to describe the human life span at the final stage, more precisely, we want to slightly expand the results of Rootzen and Zholud (2017) and find a distribution that more or less adequately describes the residual time of human life after 80 years.

Let S be some system that consists of subsystems $S^{(1)}$, $S^{(2)}$, ..., $S^{(m)}$, which work independently. Let us assume that the failure of system S occurs when at least one of the subsystems $S^{(k)}$, $k = 1, 2, \dots, m$ fails. If we denote by $\mathfrak{T}(S)$ and $\mathfrak{T}(S^{(k)})$ the uptime of system S and subsystem $S^{(k)}$ respectively, then it is clear that

$$\mathfrak{T}(S) = \min_{1 \le k \le m} \mathfrak{T}(S^{(k)}). \tag{3}$$

We will assume that each subsystem $S^{(k)}$ is given by a series-parallel system well known in reliability theory of the form

Thus, it is assumed that subsystem $S^{(k)}$ consists of series-connected blocks $B_j^{(k)}$, $j = 1, 2, \dots, n$, and each block $B_j^{(k)}$ consists of parallel-connected elements $e_{1,j}^{(k)}$, $e_{2,j}^{(k)}$, ..., $e_{k,j}^{(k)}$, the uptime of which is $\tau_{1,j}^{(k)}$, $\tau_{2,j}^{(k)}$, ..., $\tau_{k,j}^{(k)}$.

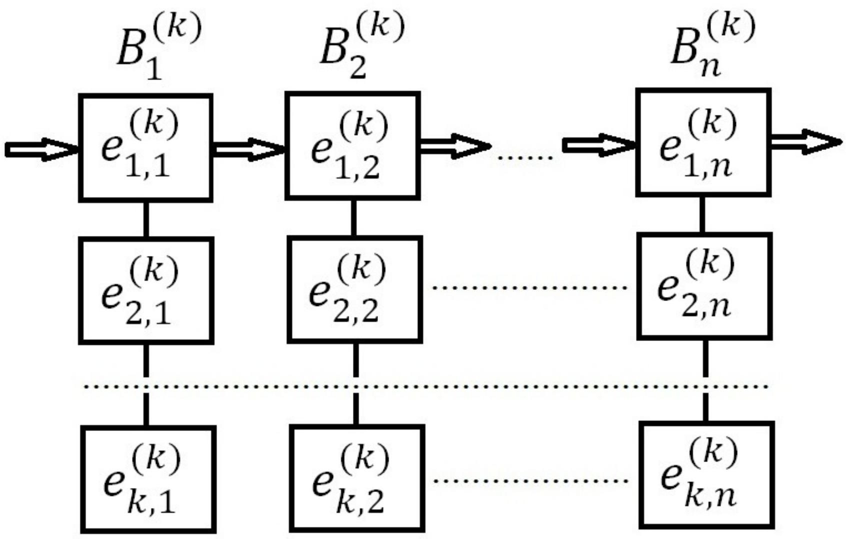

Figure 2: Series-parallel system

In addition, we impose the following condition on our model:
$\tau_{i,j}^{(k)}$, $i = 1, 2, \ldots, k$, $j = 1, 2, \ldots, n$ are independent identically distributed random variables that have an exponential distribution with parameter λ_k

$$\mathbf{P}(\tau_{i,j}^{(k)} < x) = 1 - \exp(-\lambda_k x), \quad x > 0.$$

Of course, one could assume that parameter n also depends on $k(n = n^{(k)})$. But to simplify the notation, we omit this dependence.

If we use the terminology of reliability theory, then we will consider the following two cases:

(i) loaded ("hot") reserve

and

(ii) unloaded ("cold") reserve.

Here, element $e_{1,j}^{(k)}$ is called the main element of block $B_j^{(k)}$, and $e_{2,j}^{(k)}, \ldots, e_{k,j}^{(k)}$ are reserve elements.

Let's start from case (i). The reserve elements, like the main one, are in the on state, and the moment of their connection instead of the main one does not affect their reliability. Of course, the reserve element may fail earlier than the main one (see Gnedenko, Belyaev, and Solovyev (1969), ch. 5, § 2).

In fact, from Fig.2 it follows that the failure of subsystem $S^{(k)}$ means the failure of at least one block $B_j^{(k)}$, $j = 1, 2, \ldots, n$. And block $B_j^{(k)}$ will fail when all the elements of this block fail. Then

$$\mathfrak{T}(S^{(k)}) = \min_{1 \le j \le n} \mathfrak{T}(B_j^{(k)}), \quad \mathfrak{T}(B_j^{(k)}) = \max_{1 \le i \le k} \tau_{i,j}^{(k)}, \tag{4}$$

where $\mathfrak{T}(B_j^{(k)})$ is the time to failure of block $B_j^{(k)}$.

In our conditions we have $\forall j = 1, 2, \ldots, n$

$$G_j^{(k)}(x) = \mathbf{P}(\mathfrak{T}(B_j^{(k)}) < x) = (1 - \exp(-\lambda_k x))^k, \quad x > 0. \tag{5}$$

From equality (5) we have for fixed $k \ge 1, \quad x > 0, \quad t > 0$

$$\lim_{t \to 0} \frac{G_j^{(k)}(tx)}{G_j^{(k)}(t)} = \lim_{t \to 0} \frac{(1 - \exp(-\lambda_k tx))^k}{(1 - \exp(-\lambda_k t))^k} = \lim_{t \to 0} \frac{(\lambda_k tx)^k}{(\lambda_k t)^k} = x^k. \tag{6}$$

Equality (6) means (see Leadbetter *et al.* (1983), Ch. 1, §8, Matsak (2014), Sec.2, §8), that the distribution function $G_j^{(k)}(x)$ belongs to the min-domain of attraction type 2 (Weibull distribution law):

$$\Psi_k^*(x) = \begin{cases} 0, & \text{for} \quad x \le 0, \\ 1 - \exp(-x^k), & \text{for} \quad x > 0, \end{cases} \tag{7}$$

i.e., at $n \to \infty$.

$$\mathbf{P}\left(a_n^{(k)}\left(\mathfrak{T}\left(S^{(k)}\right) - b_n^{(k)}\right) < x\right) \approx \mathbf{P}\left(a_n^{(k)}\left(\min_{1\le j\le n} \mathfrak{T}\left(B_j^{(k)}\right) - b_n^{(k)}\right) < x\right) \approx \Psi_k^*(x), \tag{8}$$

where

$$a_n^{(k)} = (\gamma_n^{(k)})^{-1},$$
$$G_j^{(k)}(\gamma_n^{(k)}) = 1 - 1/n,$$
$$b_n^{(k)} = \inf(y > 0 : G_j^{(k)}(y) > 0) = 0.$$

Approximate equality (8) can be rewritten as

$$\mathbf{P}\left(\mathfrak{T}\left(S^{(k)}\right) \ge x\right) \approx \exp\left(-\beta^{(k)} x^k\right), \quad \beta^{(k)} > 0, \quad x > 0.$$

Further, taking into account equality (3), we obtain

$$\mathbf{P}(\mathfrak{T}(S) \ge x) = \prod_{k=1}^{m} \mathbf{P}\left(\mathfrak{T}\left(S^{(k)}\right) \ge x\right) \approx \exp\left(-\sum_{k=1}^{m} \beta^{(k)} x^k\right).$$

And therefore

$$\mathbf{P}(\mathfrak{T}(S) < x) \approx W^*(\overline{\beta}, x) = 1 - \exp\left(-\sum_{k=1}^{m} \beta^{(k)} x^k\right), \tag{9}$$

$$\beta^{(k)} \ge 0, k = \overline{1, m}, \quad x \ge 0.$$

Case (ii). It is assumed that the main element $e_{1,j}^{(k)}$ of block $B_j^{(k)}$ is in the on state, and its reserve elements $e_{2,j}^{(k)}, \ldots, e_{k,j}^{(k)}$ are off. It is assumed that until they are turned on instead of the main element, they cannot have a failure.
Then

$$\mathfrak{T}(B_j^{(k)}) = \sum_{i=1}^{k} \tau_{i,j}^{(k)}.$$

As is well known, this implies the following equality

$$G_j^{(k)}(x) = \mathbf{P}(\mathfrak{T}(B_j^{(k)}) < x) = 1 - \sum_{i=0}^{k-1} \frac{(\lambda_k x)^i}{i!} \exp(-\lambda_k x), \quad x > 0. \tag{10}$$

Next, we use the following estimates:

$$\frac{(\lambda_k t)^k}{k!} = \frac{1}{1 + d_k} \sum_{i=k}^{\infty} \frac{(\lambda_k t)^i}{i!},$$

where

$$d_k = \sum_{i=k+1}^{\infty} \frac{k!}{i!} (\lambda_k t)^{i-k} \le \frac{\lambda_k t/k}{1 - \lambda_k t/k} \quad \to 0, \quad t \to 0$$

(see Leadbetter *et al.* (1983), Ch.1, §7, Example 1.7.14).
Then for $t \to 0$

$$1 - \sum_{i=0}^{k-1} \frac{(\lambda_k t)^i}{i!} \exp(-\lambda_k t) = \sum_{i=k}^{\infty} \frac{(\lambda_k t)^i}{i!} \exp(-\lambda_k t) = (1 + o(1)) \frac{(\lambda_k t)^k}{k!} \exp(-\lambda_k t). \quad (11)$$

From equalities (10),(11) we obtain for a fixed $k \geq 1, \quad x > 0, \quad t > 0$

$$\lim_{t \to 0} \frac{G_j^{(k)}(tx)}{G_j^{(k)}(t)} = \lim_{t \to 0} (1 + o(1)) \frac{(\lambda_k tx)^k}{(\lambda_k t)^k} \exp(-\lambda_k tx + \lambda_k t) = \lim_{t \to 0} \frac{(\lambda_k tx)^k}{(\lambda_k t)^k} = x^k. \quad (12)$$

That is, we have obtained an asymptotics similar to relation (6). Thus, for case (ii) the same as in case (i), equality (9) is true.

Remark 1. *It can be assumed that the basic "subsystems" on which the functioning of the human body is based somewhere, for example, after 80-90 years, already have "limited redundancy" (the numbers m and k in our model are very small). Therefore, it is natural to consider the above mathematical model as some rough approximation to a real complex biological model of the functioning of the human body at the final stage of life.*

The distribution, which is given by formula (9), we will call the W^* - distribution. In fact, this is the distribution of r.v.

$$\mathbb{W}^* = \min_{1 \leq k \leq m} \mathbb{W}_k,$$

where $\mathbb{W}_k$ are independent Weibull variables with shape k.

Further, we use this distribution in the analysis of real data on the residual time of human life after 80 years. At the same time, in formula (9) we restrict ourselves to small values of m, $(m = 2, 3)$.

3. Statistical data analysis for residual lifetime

Denote by $\mathfrak{T}_{t_0}$ the residual time of a human life that has reached the age of t_0.
And let's put forward an assumption : for sufficiently large t_0 the following hypothesis

$$H_0 = \{\mathfrak{T}_{t_0} \quad \text{a random variable that has an distribution } F_0\}$$

is true.

Our efforts in the future will be mainly directed to testing the hypothesis H_0 for the cases when we choose the distribution W^* (9), or the Gompertz (G) distribution, or the Weibull distribution as F_0.

Recall that the Weibull distribution is one of the most popular distributions in reliability theory. Its distribution function is given by:

$$F_W(x) = 1 - \exp(-\lambda x^{\alpha}), \quad \lambda \geq 0, \quad \alpha \geq 0, \quad x \geq 0.$$

In addition, we will consider one variant of the logistic distribution (in Thatcher *et al.* (1998) it is called Kannisto (LK) distribution). The Kannisto distribution has the following mortality intensities (see Thatcher *et al.* (1998)):

$$r(x) = \lambda + \frac{a \exp(bx)}{1 + a \exp(bx)}, \quad \lambda \geq 0, \quad a \geq 0, \quad b > 0, \quad x \geq 0,$$

and distribution function

$$F_{LK}(x) = 1 - \exp(-\lambda x) \left(\frac{1 + a \exp(bx)}{1 + a} \right)^{-1/b}, \quad x \geq 0. \quad (13)$$

This distribution in Thatcher *et al.* (1998) as one of the best for the analysis of old human age.

Suppose that, in a certain region, at age intervals $(t_0, t_0 + a_1)$, $(t_0 + a_1, t_0 + a_2), \ldots$, $(t_0 + a_{r-1}, t_0 + a_r)$, n_1, n_2, ..., n_r people died, respectively, $N = N(t_0) = n_1 + n_2 + \ldots + n_r$ is the number of all people in this region who have reached the age of t_0.

We introduce statistics χ^2 n the case of distribution (W^*):

$$\chi^2(\overline{\beta}) = \sum_{i=1}^{r} \frac{(n_i - Np_i(\overline{\beta}))^2}{Np_i(\overline{\beta})},$$

where $p_i(\overline{\beta}) = F_{W^*}(\overline{\beta}, a_i) - F_{W^*}(\overline{\beta}, a_{i-1})$; F_{W^*} is given by the equality (9); $a_0 = 0$.

The statistic χ^2 for the case of other distributions is introduced in a similar way.

Following the H. Cramer (see Cramér (1946), Ch.30, §3) to test hypothesis H_0 we use the minimum statistics method χ^2:

$$\chi^2(\overline{\beta}) \to \min_{\beta_i \geq 0, i=\overline{1,m}} \chi^2(\bar{\beta}). \tag{14}$$

True, unlike the H. Cramer, who reduces the problem of minimizing (14) to solving a system of nonlinear equations, we directly solve problem (14). In this case, we used the functions of the Optimization Toolbox package of the MATLAB system. For example, we used function *fminsearch*, but condition $\beta_i \geq 0, i = \overline{1,m}$ was additionally taken into account. It should also be noted that it is very important to choose a good starting point from which the minimization algorithm starts.

Hypothesis H_0 is tested according to the traditional scheme:
if $\chi^2_{\min} > x_{kr}$, then H_0 is not accepted, otherwise we consider that the data do not contradict hypothesis H_0, where $x_{kr} = U_{0.95}$ is the quantile of level 0.95 of the distribution χ^2 with $(r - m - 1)$ degrees of freedom, and $\chi^2_{\min}$ - is the minimum value of statistics χ^2 obtained from task (14).

In addition, for ungrouped samples from base IDL the goodness of fit criteria Kolmogorov-Smirnov, Cramer-Von Mises ω^2, Anderson-Darling from the statistical package STATGRAPHICS Plus were used, well known in statistics.

Example 1 Data from base (IDL)[2].

Complete IDL database file was used for the analysis. Not all countries in this dataset contain information on semi-super centenarians (aged 105-109). But for the countries used in this study, all the data are available for both semi-super centenarians (aged 105-109) and super centenarians (aged 110+).

From the base, men and women were selected separately, who satisfied the condition:

(A) born before 1902 inclusive and lived at least 105 years.

The selection of data sets prior to birth year 1902 was intended to ensure representative samples without censoring where possible.

To analyze the data on people, the advancing powers with the largest samples were selected, which satisfied the mind (A). At the time of the analysis, we obtained the following results: France (N=378), USA (N=191), England (N=149), Germany (N=108), Belgium (N=58) (N - sample size). And the hypothesis was tested that the residual time of a man's life after 105 years has an exponential distribution (that is, hypothesis H_0 was tested at $F_0 = W^*$, $t_0 = 105$, $m = 1$).

[2]International database on longevity (IDL).
URL: https://www.supercentenarians.org/en/data-and-metadata/

As it turned out, criteria Kolmogorov-Smirnov, Cramer-Von Mises ω^2, Anderson-Darling and χ^2 from package STATGRAPHICS Plus confirmed this hypothesis at a standard level of 0.05 for all the countries listed above. This naturally implies the assumption that in the general case for men the hypothesis about the exponential distribution for $t_0 = 105$ is true.

True, it should be noted that in this case, the homogeneity of the samples is not always satisfied. For example, Kolmogorov-Smirnov criterion rejected the hypothesis about the same distribution of samples of men from France and the United States, but accepted the same hypothesis for France and Germany, as well as for France and Belgium.

Turning to the analysis of data on women, we immediately note that the situation here is a little more complicated. The following table contains the results of calculating the minimum values of χ^2 statistics for women in France, Germany and a pooled sample of BCN (Belgium, Canada and Norway).

The calculation of the best parameters for distributions $W*$, G, K and W (for which the observed value of Chi^2 was minimal) was performed using function $[x, fval] = fminsearch()$ in package MATLAB. The number of intervals was adjusted as follows: if the number of elements in one of the extreme intervals was less than 10, then this and subsequent intervals were combined into one.

The MATLAB code for calculating values for France is given in Appendix 1.

Table 1: Data from base IDL, women born in 1902 or earlier

Country	France	Germany	BCN
t_0	105	105	105
$N(t_0)$	4166	829	867
$\chi^2(W^*)$	17.91	9.55	8.04
$\chi^2(G)$	15.78	9.23	8.04
$\chi^2(LK)$	15.77	8.37	8.04
$\chi^2(W)$	33.97	11.98	9.13
$r - m - 1$	16	8	8
$U_{0.95}$	26.3	15.5	15.5

Notations in Table 1:

t_0 is the age from which the residual life time is considered;
$N(t_0)$ is the number of people who have reached age t_0;
$\chi^2(W^*)$ is the minimum value of the statistics χ^2 for distribution W^* at $m = 2$;
$\chi^2(G)$ is the minimum value of the statistics χ^2 for the Gompertz(G) distribution (1);
$\chi^2(LK)$ is the minimum value of the statistics χ^2 for the Kannisto LK distribution ;
$\chi^2(W)$ is the minimum value of the statistics χ^2 for the Weibull distribution;
$r - m - 1$ is the number of degrees of freedom, r is the number of grouping intervals, m is the number of unknown parameters of the respective distribution;
$U_{0.95}$ - quantile level 0,95 distribution χ^2 with $r - m - 1$ degrees of freedom.

Comments on Table 1.

(i) According to Table 1 criterion χ^2 accepted hypothesis H_0 for the distribution of W^* at $t_0 = 105$, $m = 2$ for women in France, Germany and a pooled sample of BCN.

A similar situation persists for the Gompertz (G) and Kannisto (LK) distributions.

(ii) Thus criterion χ^2 itself accepted the hypothesis that residual life after 105 years is described by the Weibull distribution for women in Germany and the pooled sample BCN, but rejected this hypothesis for France.

(iii) It should also be noted that for women in France and Germany, criteria Kolmogorov-Smirnov, Cramer-Von Mises ω^2, Anderson-Darling from the STATGRAPHICS Plus package rejected the Weibull distribution.

(iv) If we use the terminology of reliability theory, then we can say (probably not entirely accurate) that for men, the system for protecting the functioning of the body reaches the “zero level of redundancy” somewhere around 105 years old, and for women - 5 years later, that is, at 110 years old .

Example 2. Mortality data for centenarians in the Netherlands.

Relevant data was selected from article K.Aarssen, L.de Haan (Karin and de Haan (1994), p.266, Tabl.2). he authors provide separately for men and women born in 1877-1881 data on mortality of centenarians at the final stage of life.

The following two tables contain the results of the analysis of these data. More precisely, we present the calculated minimum values of the statistics χ^2 for the residual life time after reaching a certain age of $t_0 \approx 90$ years.

Table 2: Netherlands, men, born 1877-1881

Year of birth	1877	1878	1879	1880	1881	1877-1881 (total)
t_0	94	93	92	91	90	94
$N(t_0)$	815	1092	1575	2062	2731	4131
$\chi^2(W^*)$	4.49	14.13	7.12	10.1	7.02	3.52
$\chi^2(G)$	4.41	14.58	8.27	10.87	11.33	5.39
$\chi^2(LK)$	4.13	14.72	7.25	10.17	6.41	3.72
$\chi^2(W)$	6.63	17.94	8.58	15.1	16.9	10.67
$r-m-1$	7	10	10	11	11	10
$U_{0.95}$	14.1	18.3	18.3	19.67	19.67	18.3

Table 3: Netherlands, women, born 1877-1881

Year of birth	1877	1878	1879	1880	1881	1877-1881 (total)
t_0	94	93	92	91	90	94
$N(t_0)$	1097	1556	2266	2939	3943	6260
$\chi^2(W^*)$	11.04	9.2	19.21	14.66	9.89	14.52
$\chi^2(G)$	11.0	8.97	15.97	14.87	15.28	13.38
$\chi^2(LK)$	11.15	10.45	19.03	16.53	11.35	13.66
$\chi^2(W)$	12.54	16.59	34.29	33.95	23.88	35.93
$r-m-1$	9	10	10	12	13	11
$U_{0.95}$	16.9	18.3	18.3	21.03	22.4	19.67

The MATLAB code for calculating values for women is given in Appendix 3.

Comments on tables 2, 3.

(i) From Tables 2 and 3 we conclude that for distributions W^*, Gompertz(G) and Kannisto(K) hypothesis H_0 is true in almost all cases.

(ii) The Weibull distribution is inconsistent with the data for women born 1879-1881, as well as for the total data for women born 1877-1881.

(iii) W^* distribution at $m = 2$ in all cases (except for women born in 1879) does not contradict the data (replacing $m = 2$ gives $m = 3$ gives $\chi^2(W^*) = 11.33$, for the case of women born in 1879, i.e. hypothesis H_0 is accepted).

(iv) Thus, the data from Karin and de Haan (1994) do not contradict the hypothesis that distributions W^*, Gompertz (G) and Kannisto (LK) adequately describe the residual time of human life after $t_0 = 90$ years.

Example 3. Mortality data in Germany, born 1895-1899.

The data were selected from the open online database GFGI[3]. The database contains a large number of records with information about the birth, death, etc. of the German population since 1600. For the analysis, only the data "date of birth" and "date of death" were used. Records that did not contain the exact date of birth or death (day, month, year) were excluded from analysis. Using Excel function $= DATEDIF(BDATE; DDATE; "y")$, the number of complete years was calculated for each record. Using filters in the database, we obtained samples that contained information about people born in 1895-1899.

Similarly, as above, the minimum values of statistics χ^2 for residual life time were calculated separately for men and women after 80 years.

Since the situation here is a bit more complicated than in the previous cases, the parameter $m = 3$ was chosen for the distribution W^*, and the Gompertz (G) distribution remained unchanged (the transition to a more general Gompertz-Makeham (GM) distribution did not significantly change the situation).

In the following tables 4 and 5 the calculation results are given. The corresponding grouped data, according to which tables 4, 5 were built, are given at the end of the article (Appendix 2, Table 9).

Table 4: Germany, data from the GFGI database, men, born 1895 - 1899

Year of birth	1895	1896	1897	1898	1899	1895-1899 (total)
t_0	80	80	80	80	80	80
$N(t_0)$	634	676	784	811	894	3799
$\chi^2(W^*)$	11.4	16.39	17.85	23.93	15.3	39.68
$\chi^2(G)$	11.46	16.56	22.80	24.45	15.21	45.18
$\chi^2(LK)$	13.10	19.93	26.00	24.09	14.15	40.46
$\chi^2(W)$	30.0	36.99	57.45	37.25	32.5	118.81
$r - m - 1$	15	16	16	13	15	19
$U_{0.95}$	24.99	26.3	26.3	22.36	24.99	30.14

Table 5: Germany, data from the GFGI database, women, born 1895 - 1899

Year of birth	1895	1896	1897	1898	1899	1895-1899 (total)
t_0	80	80	80	80	80	80
$N(t_0)$	1039	1135	1181	1263	1368	5986
$\chi^2(W^*)$	23.66	11.4	20.7	25.57	15.16	22.6
$\chi^2(G)$	27.60	14.49	18.67	26.80	14.99	21.93
$\chi^2(LK)$	28.59	12.59	22.06	27.03	15.55	23.12
$\chi^2(W)$	34.27	38.85	70.36	90.94	78.74	236.85
$r - m - 1$	17	18	19	19	20	22
$U_{0.95}$	27.58	28.87	30.1	30.1	31.4	33.9

Comments to tables 4, 5.

(i) From tables 4, 5 From tables 4, 5 it is clear that for men and women in Germany, the final life expectancy after 80 years is much better described by the W^* distribution than by the Weibull distribution. More precisely, distribution W^* at $m = 3$ is accepted for women

[3]Germany, Find a Grave™ Index, 1600s-Current (GFGI). [database on-line]. Lehi, UT, USA: Ancestry.com Operations, Inc., 2012.
URL: https://www.ancestry.com/search/collections/60533/

in all cases (table 5), and in most cases for men (table 4). At the same time, the Weibull distribution deviates in all cases. For cases of Gompertz (G) and Kannisto (LK) we have values of the χ^2 statistic close to the $\chi^2(W^*)$ statistic.

(ii) If in table 4 in column "1895-1899 (total)" we consider the total data without 1898 year of birth, then we get $\chi^2(W^*) = 24.79$, that is, hypothesis H_0 for W^* distribution is accepted. This indicates certain differences (heterogeneity) in the data for the 1898 year of birth, in contrast to other years.

(iii) Note that when calculating tables 2-5 in the case of the Kannisto(LK) you can choose the parameter $\lambda = 0$. This will not significantly change the value of the χ^2 statistic.

Remark 2. *If, in the conditions of example 3, we choose $t_0 = 75$, that is, consider the residual time of human life after 75 years, then the situation will become much more complicated. So hypothesis H_0 for the distribution of W^* at $t_0 = 75, m = 3$ at a significance level of $\alpha = 0.05$ is accepted somewhere in half of the cases.*

Remark 3. *On the data from example 3, preliminary comparisons of the Gamma distribution and some others with the W^* distribution were also carried out. As it turned out, their performance for the χ^2 criterion was worse than that of the W^* distribution.*

Example 4. Mortality data in Norway, born 1881-1885.

The data were selected from the online database HMD[4]. For the study, we used dataset `Deaths_1x1`.

Similarly, the minimum values of statistics χ^2 for residual life time were calculated separately for men and women after 80 years. In this case, for the W^* distribution the parameter $m = 3$ was chosen. The following tables 6 and 7 give the calculation results. The corresponding grouped data, according to which tables 6, 7 were built, are given at the end of the work (Appendix 2, Table 10).

Table 6: Norway, data from the HMD database, men, born 1881 - 1885

Year of birth	1881	1882	1883	1884	1885	1881-1885 total)
t_0	80	80	80	80	80	80
$N(t_0)$	4780	5179	5145	5289	5492	25884
$\chi^2(W^*)$	6.66	11.07	16.58	6.97	15.49	7.84
$\chi^2(G)$	7.78	12.78	14.20	8.36	17.11	14.69
$\chi^2(LK)$	8.65	12.75	21.81	6.84	15.37	9.70
$\chi^2(W)$	174.35	194.41	288.94	147.89	169.93	842.22
$r - m - 1$	20	20	20	20	20	20
$U_{0.95}$	31.4	31.4	31.4	31.4	31.4	31.4

Comments to the tables 6, 7.

(i) From Tables 6 and 7 we draw a conclusion: for men and women in Norway, the final life expectancy after 80 years is adequately described by distribution W^* (at a significance level of 0.05). The same applies to the Kannisto (LK) distribution. The Gompertz (G) distribution is also accepted almost always, with the exception of summary data for women. The Weibull distribution is rejected in all cases.

(ii) It should be noted that the results for men and women in Norway were not successful for other countries from base HMD. Perhaps this is due both to some heterogeneity in real data and to some changes made to them during their fixation and initial processing.

[4]Human Mortality Database (HMD).
URL: https://mortality.org/Home/Index

Table 7: Norway, data from the HMD database, women, born 1881 - 1885

Year of birth	1881	1882	1883	1884	1885	1881-1885 (total)
t_0	80	80	80	80	80	80
$N(t_0)$	6498	6927	7137	7360	7480	35402
$\chi^2(W^*)$	6.30	13.49	20.04	9.45	8.10	12.13
$\chi^2(G)$	14.94	18.93	30.99	18.52	19.72	57.12
$\chi^2(LK)$	7.72	13.39	21.10	10.24	10.72	16.41
$\chi^2(W)$	184.54	226.82	265.65	230.91	199.53	1.05e+03
$r-m-1$	20	20	20	20	20	20
$U_{0.95}$	31.4	31.4	31.4	31.4	31.4	31.4

Example 5. Total deaths in Europe after 80 years (11 countries, cohorts born 1871-1880, table 3 from book Thatcher *et al.* (1998)).

In our opinion, the total data on mortality (11 countries + 10 cohorts) over the interval (80, 120) years can hardly be considered homogeneous. For example, for this data, the calculated minimum values of statistic χ^2 for W^* distribution were:

$$\chi^2(W^*) = 124.70 \quad \text{for men,}$$

and

$$\chi^2(W^*) = 97.44 \quad \text{for women,}$$

which is several times higher than the critical values at a significance level of $\alpha = 0.05$. The minimum values of the χ^2 statistic for the Kannisto (LK), Gompertz (G) and Weibull (W) distributions were even larger.

It can be hoped that the situation will improve when considering the residual time of human life somewhere after 100 years. Therefore, in the following table, the interval (99, 120) years is considered (a similar interval was also chosen in work Thatcher *et al.* (1998), table 6.4). The notation is similar to that introduced above.

Table 8: Analysis of data on mortality of centenarians in Europe (residual life after 99 years, 11 countries, cohorts born 1871-1880)

sex	men	women	together
t_0	99	99	99
$N(t_0)$	12967	44642	57609
$\chi^2(W^*)$	17.27	15.77	24.71
$\chi^2(G)$	18.03	15.57	24.48
$\chi^2(LK)$	16.86	19.16	26.62
$\chi^2(W)$	11.86	76.0	76.92
$r-m-1$	8	8	8
$U_{0.95}$	15.5	15.5	15.5

Interestingly, the minimum values of statistics $\chi^2(W^*)$, $\chi^2(G)$ and $\chi^2(LK)$ in the above table 8 are significantly lower (especially for men) than in the corresponding table 6.4 of Thatcher *et al.* (1998). Perhaps this is explained by different methods for calculating the minimum χ^2.

4. Conclusions

In this work, the task was to find a simple distribution that can be used to describe the residual time of a human life after 80 years with a standard reliability of 0.95 for applied statistics.

In our opinion, the above preliminary calculations on real data show that for this we can try to use the W^* and Kannisto (LK) distributions, and possibly the Gompertz (G) distributions for fairly homogeneous data.

From the calculations in Table 8, it can be assumed that for summaries of data for several countries and cohorts, these distributions can probably be used somewhere after 100 years.

Of course, the results presented in tables 1 - 8 are not yet sufficient to establish the validity of hypothesis H_0 for the W_m^*, Gompertz (G) and Kannisto (LK) distributions.

It is clear that further tests on real data are needed for final conclusions.

References

Bowers NL, Gerber HU, Hickman JC, Jones DA, Nesbitt CJ (1997). *Actuarial Mathematics*. Second edition. The Society Of Actuaries.

Comfort A (1964). *Ageing: The Biology of Senescence*. Routledge & Kegan Paul.

Cramér H (1946). *Mathematical Methods of Statistics*. University Press.

Economos A (1983). "Rate of Aging, Rate of Dying and the Mechanism of Mortality." *Archives of Gerontology and Geriatrics*, **1**(1), 3–27. doi:10.1016/0167-4943(82)90003-6.

Feller W (1957). *An Introduction to Probability Theory and Its Applications*. 1 edition. John Wiley and Sons, New York, London.

Gavrilov L, Gavrilova N (1991). *The Biology of Life Span: A Quantitative Approach*. Harwood Academic Publisher, London.

Gnedenko B (1943). "Sur la Distribution Limit du Terme Maximum d'Une Serie Aleatoire." *Annals of Mathematics*, **44**(3), 423–453. doi:10.1080/08898489409525379.

Gnedenko B, Belyaev Y, Solovyev A (1969). *Mathematical Methods of Reliability Theory*. Academic Press, New York, London.

Gompertz B (1825). "On the Nature of the Function Expressive of the Law of Human Mortality and on a New Mode of Determining Life Contingencies." *Philosophical Transactions of the Royal Society*, **115**(115), 513–585.

Greenwood M, Irwin J (1939). "The Biostatistics of Senility." *Human Biology*, **11**, 1–23.

Karin A, de Haan L (1994). "On the Maximal Life Span of Humans." *Mathematical Population Studies*, **4**(4), 259–81. doi:10.1080/08898489409525379.

Leadbetter M, Lindgren G, Rootzen H (1983). *Extremes and Related Properties of Random Sequences and Processes*. Springer, New York. doi:10.1007/978-1-4612-5449-2.

Makeham WM (1860). "On the Law of Mortality and the Construction of Annuity Tables." *The Assurance Magazine, and Journal of the Institute of Actuaries*, **8**(6), 301–310.

Matsak I (2014). *Elements of the Theory of Extreme Values*. CPU "COMPRINT", Kyiv (In Ukrainian).

Modig K, Andersson T, Vaupel J, Rau R, Ahlbom A (2017). "How Long Do Centenarians Survive? Life Expectancy and Maximum Lifespan." *Journal of Internal Medicine*, **282**(2), 156–163. doi:10.1111/joim.12627.

Rootzen H, Zholud D (2017). "Human Life Is Unlimited – but Short." *Extremes*, **20**, 713–728. doi:10.1007/s10687-017-0305-5.

Rosset E (1981). *Human Longevity*. Progress, Moscow.

Thatcher A, Kannisto V, Vaupel J (1998). *The Force of Mortality at Ages 80 to 120*. Syddansk Universitetsforlag. Odense Monographs on Population Aging No.5.

Thatcher R (1999). "The Long-term Pattern of Adult Mortality and the Highest Attained Age." *Journal of the Royal Statistical Society. Series A (Statistics in Society)*, **162**(1), 5–43. doi:10.1111/1467-985x.00119.

Appendix 1

Calculations of data for French women from the IDL database (105+) in Table 1.

1) Distribution (W^*) is given by formula (9).

Optimal parameter values:

```
(beta_i =(x_i)^2) -> x=[0.0385, 0.0006];
```

The value of the χ^2 statistic at the optimal point:

```
 [xmin, fval]=fminsearch('rtl_fra_1902wf', [0.0489, 0.0014])
xmin =
     0.0385      0.0006
fval =
    17.9118
```

A function that calculates the values of the χ^2 statistic for distribution (W^*):

```
function rtl_fra_1902w=rtl_fra_1902wf(x)
fi=[689   586   502   427   357 314 268 220   168    132    114   102    75
51     52    35     31    24   19];
xn=[0     115   230   345   460 575 690 805   920   1035   1150   1265 1380
1495 1610   1725   1840 1955   2070];
xv=[115   230   345   460   575 690 805 920   1035 1150   1265   1380 1495
1610 1725   1840   1955 2070   2500];
n=sum(fi);
s2=1-exp(-x(1)^2 .*xv -x(2)^2 .*(xv.^2));
s1=1-exp(-x(1)^2 .*xn -x(2)^2 .*(xn.^2));
npi=(s2-s1).*n;
rtl_fra_1902w= sum(((fi-npi).^2)./npi);
```

2) Gompertz distribution (G) is given by formula (1) at $\lambda = 0$.

Optimal parameter values:

```
(alpha=(x_2)^2 , beta=(x_1)^2) -> x=[2.0774, 0.0186];
```

The value of the χ^2 statistic at the optimal point:

```
[xmin, fval]=fminsearch('g_fra_1902wf' ,[0.0489 ,  0.0014])
xmin =
     2.0774      0.0186
fval =
    15.7790
```

A function that calculates the values of the χ^2 statistic for Gompertz distribution (G):

```
function g_fra_1902w = g_fra_1902wf(x)
fi=[689    586    502     427   357   314   268   220    168    132     114
102    75         51     52       35       31      24     19];
xn=[0       115   230   345   460 575 690 805   920   1035   1150   1265 1380
1495 1610   1725   1840 1955   2070];
xv=[115   230   345   460   575 690 805 920   1035 1150   1265   1380 1495
1610 1725   1840   1955 2070   2500];
n=sum(fi);
s2=1-exp(-(x(1)^2)*(exp(x(2)^2 .*xv)-1));
s1=1-exp(-(x(1)^2)*(exp(x(2)^2 .*xn)-1));
npi=(s2-s1).*n;
g_fra_1902w = sum(((fi-npi).^2)./npi);
```

3) Konnisto distribution (LK) is given by formula (13).

Optimal parameter values:

```
(lambda=(x_1)^2), a=(x_2)^2, b=(x_3)^2) -> x=[0.0033, 0.0386, 0.0187];
```

The value of the χ^2 statistic at the optimal point:

```
[xmin, fval]=fminsearch('LKfraf_1902f' ,[0, 0.0489, 0.0014])
xmin =
    0.0033      0.0386      0.0187
fval =
   15.7714
```

A function that calculates the values of the χ^2 statistic for Konnisto distribution (LK):

```
function LKfraf_1902=LKfraf_1902f(x)
fi=[689     586   502     427   357    314   268   220    168    132    114
102     75      51     52     35      31     24        19];
xn=[ 0     115   230   345   460  575  690  805   920   1035   1150   1265  1380
1495  1610   1725   1840  1955   2070];
xv=[115   230   345   460   575  690  805  920   1035  1150   1265   1380  1495
1610  1725   1840   1955  2070   2500];
n=sum(fi);
s21=exp(-x(1)^2 .*xv);
s22= ((1+x(2)^2 *exp(x(3)^2 .*xv))/(1+x(2)^2)).^(-1/x(3)^2);
s2=1-s21 .*s22;
s11=exp(-x(1)^2 .*xn);
s12= ((1+x(2)^2 *exp(x(3)^2 .*xn))/(1+x(2)^2)).^(-1/x(3)^2);
s1=1-s11 .*s12;
npi=(s2-s1).*n;
LKfraf_1902 = sum(((fi-npi).^2)./npi);
```

4) Weibull distribution (W).

Optimal parameter values:

```
(lambda=(x_1)^2 , alpha=(x_2)^2) -> x=[0.0312, 1.0866];
```

The value of the χ^2 statistic at the optimal point:

```
[xmin, fval]=fminsearch('wey_fra_1902wf',[0.0489, 0.0014])
xmin =
    0.0312      1.0866
fval =
   33.9747
```

A function that calculates the values of the χ^2 statistic for Weibull distribution (W):

```
function wey_fra_1902w=wey_fra_1902wf(x)
fi=[689   586   502   427   357  314  268  220   168    132    114   102    75
51     52    35     31    24   19];
xn=[0     115   230   345   460  575  690  805   920   1035   1150   1265  1380
1495  1610   1725   1840  1955   2070];
xv=[115   230   345   460   575  690  805  920   1035  1150   1265   1380  1495
1610  1725   1840   1955  2070   2500];
n=sum(fi);
n=sum(fi);
s2=1-exp(-x(1)^2.*(xv.^x(2)));
s1=1-exp(-x(1)^2.*(xn.^x(2)));
```

```
npi=(s2-s1).*n;
wey_fra_1902w= sum(((fi-npi).^2)./npi);
```

Appendix 2

Table 9: Germany, data from the GFGI database, born 1895-1899

Year of birth	1895		1896		1897		1898		1899	
Age	m	w	m	w	m	w	m	w	m	w
75	59	53	63	56	42	52	63	59	78	54
76	59	71	60	45	62	65	81	61	107	65
77	51	76	63	54	81	68	79	78	86	73
78	43	68	66	80	87	82	92	60	86	85
79	53	56	70	87	83	107	84	59	74	93
80	70	49	59	69	91	79	74	99	70	85
81	56	72	52	72	67	78	64	80	78	92
82	61	91	62	79	70	67	68	87	85	85
83	55	74	54	91	72	86	89	93	73	85
84	48	69	52	90	55	88	71	64	73	88
85	50	75	53	91	59	74	67	91	79	103
86	51	88	35	76	38	77	53	84	63	91
87	43	54	43	64	44	83	59	85	64	82
88	24	72	35	67	49	71	34	76	55	76
89	34	64	49	76	40	70	38	95	44	85
90	27	53	31	59	42	59	43	65	46	87
91	23	42	29	53	40	48	51	69	30	58
92	15	42	30	48	20	37	25	51	36	67
93	16	43	22	51	25	43	21	48	30	63
94	19	33	9	33	15	49	14	37	13	42
95	13	31	15	28	18	37	10	32	13	45
96	10	17	14	23	10	31	6	26	13	30
97	3	20	11	15	10	33	6	15	15	32
98	8	19	5	17	4	23	8	18	4	17
99	1	7	4	8	6	12	4	16	4	14
100	1	7	3	9	6	13	1	11	1	13
101	0	5	3	2	0	10	2	12	2	12
102	4	6	2	3	2	6	1	4	1	8
103	1	4	1	5	0	2	2	2	1	4
104	0	2	3	2	0	2	0	2	0	2
105	0	0	0	2	0	1	0	1	1	2
106	0	0	0	1	1	1	0	0	0	0
107	1	0	0	1	0	1	0	0	0	0
Total 75+	899	1363	998	1457	1139	1555	1210	1580	1325	1738
Total 80+	634	1039	676	1135	784	1181	811	1263	894	1368

Table 10: Norway, data from the HMD database, born in 1881-1885

Year of birth	1881		1882		1883		1884		1885	
Age	m	w	m	w	m	w	m	w	m	w
80	492	525	540	608	524	613	533	607	553	603
81	494	553	523	619	484	586	531	600	561	599
82	470	568	487	590	480	568	500	584	550	617
83	429	553	483	579	479	564	474	595	529	632
84	403	547	424	559	437	560	471	597	478	616
85	366	518	396	526	403	540	434	565	431	587
86	335	484	402	485	388	506	369	533	381	551
87	327	436	373	463	369	513	330	517	349	502
88	291	390	323	442	320	502	309	470	328	456
89	244	363	263	384	268	430	268	407	277	418
90	219	317	220	335	226	354	224	361	236	367
91	178	261	185	295	184	297	197	320	206	317
92	136	227	142	246	145	246	157	270	166	266
93	105	193	117	203	113	206	129	216	129	218
94	86	159	96	151	89	156	107	171	87	182
95	70	120	68	111	74	131	79	156	65	150
96	46	85	48	93	60	111	58	117	50	116
97	31	66	36	84	41	77	39	84	44	92
98	25	42	22	51	24	65	28	57	18	56
99	12	30	14	31	16	35	15	43	25	53
100	9	21	10	18	12	27	14	31	10	26
101	8	16	3	22	7	18	9	24	9	20
102	3	14	4	11	6	14	8	9	6	11
103	5	5	2	10	0	10	4	12	3	11
104	0	3	1	5	1	5	3	7	0	9
105	0	2	1	4	0	3	1	8	1	3
106	0	2	0	2	0	3	1	0	0	0
107	0	2	2	1	0	1	0	2	1	3
108	0	0	0	0	0	0	0	0	0	2
109	0	0	0	0	0	1	0	0	0	1
Total 80+	8646	10465	9225	11267	9381	11403	9686	11947	9974	12182

Appendix 3

Calculations of data for Netherlands women in Table 3.

1) Distribution (W^*) is given by formula (9).

Optimal parameter values:

```
(beta_i =(x_i)^2) -> x=[0.5495, 0.1066];
```

The value of the χ^2 statistic at the optimal point:

```
rtl_wsuf(x)
ans =
    14.5548
```

A function that calculates the values of the χ^2 statistic for distribution (W^*):

```
function axiwsu=rtl_wsuf(x)
fi=[1701 1273 1038 705 476 362 249 192 110 66 35 22 15 16];
xn=[0    1    2    3   4   5   6   7   8   9  10 11 12  13];
xv=[1    2    3    4   5   6   7   8   9   10 11 12 13  50];
n=sum(fi);
s2=1-exp(-x(1)^2 .*xv -x(2)^2 .*(xv.^2));
s1=1-exp(-x(1)^2 .*xn -x(2)^2 .*(xn.^2));
npi=(s2-s1).*n;
axiwsu= sum(((fi-npi).^2)./npi);
```

2) Gompertz distribution (G) is given by formula (1) at $\lambda = 0$.

Optimal parameter values:

```
(alpha=(x_2)^2 , beta=(x_1)^2) -> x=[2.2804, 0.2429];
```

The value of the χ^2 statistic at the optimal point:

```
g_rtl_wsuf(x)
ans =
    13.3847
```

A function that calculates the values of the χ^2 statistic for Gompertz distribution (G):

```
function g_rtl_wsu=g_rtl_wsuf(x)
fi=[1701 1273 1038 705 476 362 249 192 110 66 35 22 15 16];
xn=[0    1    2    3   4   5   6   7   8   9  10 11 12  13];
xv=[1    2    3    4   5   6   7   8   9   10 11 12 13  50];
n=sum(fi);
s2=1-exp(-(x(1)^2)*(exp(x(2)^2 .*xv)-1));
s1=1-exp(-(x(1)^2)*(exp(x(2)^2 .*xn)-1));
npi=(s2-s1).*n;
g_rtl_wsu= sum(((fi-npi).^2)./npi);
```

3) Konnisto distribution (LK) is given by formula (13).

Optimal parameter values:

```
(lambda=(x_1)^2), a=(x_2)^2, b=(x_3)^2) -> x=[0.0072, 0.6627, 0.3114];
```

The value of the χ^2 statistic at the optimal point:

```
LKrtl_wsuf(x)
ans =
    13.6640
```

A function that calculates the values of the χ^2 statistic for Konnisto distribution (LK):

```
function LKrtl_wsu=LKrtl_wsuf(x)
fi=[1701 1273 1038 705 476 362 249 192 110 66 35 22 15 16];
xn=[0    1    2    3   4   5   6   7   8   9  10 11 12  13];
xv=[1    2    3    4   5   6   7   8   9   10 11 12 13  50];
n=sum(fi);
s21=exp(x(1)^2 .*xv);
s22= ((1+x(2)^2 *exp(x(3)^2 .*xv))/(1+x(2)^2)).^(-1/x(3)^2);
s2=1-s21 .*s22;
s11=exp(x(1)^2 .*xn);
s12= ((1+x(2)^2 *exp(x(3)^2 .*xn))/(1+x(2)^2)).^(-1/x(3)^2);
s1=1-s11 .*s12;
npi=(s2-s1).*n;
LKrtl_wsu = sum(((fi-npi).^2)./fi);
```

4) Weibull distribution (W).

Optimal parameter values:

```
(lambda=(x_1)^2 , alpha=(x_2)^2) -> x=[0.5511,  1.0539];
```

The value of the χ^2 statistic at the optimal point:

```
wey_wsuf(x)
ans =
    35.9332
```

A function that calculates the values of the χ^2 statistic for Weibull distribution (W):

```
function axiwsu=wey_wsuf(x)
fi=[1701 1273 1038 705 476 362 249 192 110 66 35 22 15 16];
xn=[0    1    2    3   4   5   6   7   8   9  10 11 12  13];
xv=[1    2    3    4   5   6   7   8   9   10 11 12 13  20];
n=sum(fi);
s2=1-exp(-x(1)^2.*(xv.^(x(2)^2)));
s1=1-exp(-x(1)^2.*(xn.^(x(2)^2)));
npi=(s2-s1).*n;
axiwsu= sum(((fi-npi).^2)./npi);
```

Affiliation:

Kateryna Akbash
Volodymyr Vynnychenko Central Ukrainian State University
Shevchenko street 1, Kropyvnytskyi 25006, Ukraine
E-mail: kateryna.akbash@gmail.com

Austrian Journal of Statistics http://www.ajs.or.at/
published by the Austrian Society of Statistics http://www.osg.or.at/
Volume 53 *Submitted:* 2023-11-28
2024 *Accepted:* 2024-04-03