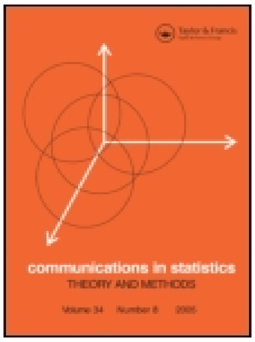

Communications in Statistics - Theory and Methods

ISSN: 0361-0926 (Print) 1532-415X (Online) Journal homepage: https://www.tandfonline.com/loi/lsta20

Goodness-of-fit tests for the Gompertz distribution

Adam Lenart & Trifon I. Missov

To cite this article: Adam Lenart & Trifon I. Missov (2016) Goodness-of-fit tests for the Gompertz distribution, Communications in Statistics - Theory and Methods, 45:10, 2920-2937, DOI: 10.1080/03610926.2014.892323

To link to this article: https://doi.org/10.1080/03610926.2014.892323

COMMUNICATIONS IN STATISTICS—THEORY AND METHODS
2016, VOL. 45, NO. 10, 2920–2937
http://dx.doi.org/10.1080/03610926.2014.892323

Taylor & Francis
Taylor & Francis Group

OPEN ACCESS

Goodness-of-fit tests for the Gompertz distribution

Adam Lenart[a,b] and Trifon I. Missov[b,c]

[a]Max Planck Odense Center on the Biodemography of Aging, Department of Epidemiology, Biostatistics and Biodemography, University of Southern Denmark, Odense, Denmark; [b]Max Planck Institute for Demographic Research, Rostock, Germany; [c]Institute of Sociology and Demography, University of Rostock, Rostock, Germany

ABSTRACT
While the Gompertz distribution is often fitted to lifespan data, testing whether the fit satisfies theoretical criteria is being neglected. Here four goodness-of-fit measures – the Anderson–Darling statistic, the correlation coefficient test, a statistic using moments, and a nested test against the generalized extreme value distributions – are discussed. Along with an application to laboratory rat data, critical values calculated by the empirical distribution of the test statistics are also presented.

ARTICLE HISTORY
Received 9 August 2012
Accepted 31 January 2014

KEYWORDS
Goodness-of-fit, Anderson-Darling statistic, Correlation coeffiicient, Moments, Extreme value distribution, Empirical critical values.

MATHEMATICS SUBJECT CLASSIFICATION
Primary: 62N03; Secondary: 62F03

1. Introduction

Goodness-of-fit tests determine if the empirical distribution of the data satisfies the assumptions of theoretical distributions. While the Gompertz distribution is routinely used as life time distribution in demography, biology, actuarial, and medical science, according to our best knowledge, no studies on goodness-of-fit tests for it have been published so far. However, the Gompertz distribution is a degenerate generalized extreme value distribution for the minima, and an abundance of goodness-of-fit tests exist in the literature for other extreme value distributions (see, e.g., Hosking, 1984).

In a landmark paper, Anderson–Darling (1952) developed the Anderson–Darling test that later Stephens (1977) analyzed in the context of extreme value distributions. Sinclair et al. (1990) modified the Anderson–Darling test to allow different weighting schemes that emphasize either the lower or the upper tail of the distributions.

CONTACT Adam Lenart alenart@health.sdu.dk Max Planck Odense Center on the Biodemography of Aging, Institute of Public Health, University of Southern Denmark, J. B. Winsløws Vej 9, 5000 Odense, Denmark.
Color versions of one or more of the figures in the article can be found online at www.tandfonline.com/lsta.

Filliben (1975) used the Pearson correlation coefficient to check the correlation between expected statistics of a theoretical distribution and sample statistics. The correlation coefficient test was the most popular in hydrology (Vogel, 1986; Kinnison, 1989) to assess the fit of extreme value distributions.

The likelihood ratio test naturally arises to account for the differences between the Gompertz and other extreme value distributions. The generalized extreme value distribution is characterized by μ, location, σ, scale, and ξ, shape, parameters. For $\xi = 0$, the generalized extreme value distribution reduces to the Gumbel, and the Gompertz distribution is a reversed and truncated Gumbel distribution with additional correlation between the maximum likelihood estimate of its parameters a and b. The different parametrization of the Gompertz distribution removes it from location-scale family of distributions.

Li and Papadopoulos (2002) proposed a goodness-of-fit test using moments. The test statistic is derived from an identity for the moments, and its values are compared to the z-values of the standard normal distribution.

This paper will first briefly describe each of these tests and apply them to the Gompertz distribution. The final sections of the paper compare the power of the tests against alternative distributions and derive critical values of them based on Monte Carlo simulation experiments. An application of the tests to laboratory rat data is also discussed.

1.1 Properties of the Gompertz distribution

The Gompertz distribution is often applied to describe the distribution of adult lifespans by demographers (e.g., Vaupel, 1986; Doblhammer, 2000; Preston et al., 2001; Willekens, 2001; Perozek, 2008) and actuaries (Benjamin et al., 1980; Willemse and Koppelaar, 2000). It is also used to fit the mortality data of birds, mammals (Finch et al., 1990; Promislow, 1991; Witten and Satzer, 1992; Finch and Pike, 1996; Ricklefs and Scheuerlein, 2002), and sometimes invertebrates (Hirsch and Peretz, 1984; Honda and Matsuo, 1992).

The Gompertz distribution has a continuous probability density function with parameters a and b,

$$f(x) = ae^{bx-\frac{a}{b}\left(e^{bx}-1\right)} \qquad a \geq 0, b > 0, \tag{1}$$

with support on $[0, \infty)$. Please see Fig. 1 for the shape of the Gompertz distribution.

Given its popularity, the Gompertz distribution is surprisingly understudied in the statistical, demographic literature. Pollard and Valkovics (1992) were the first to analyze the statistical properties of the Gompertz distribution, however their results only hold asymptotically when $a \to 0$. Exact moments of the Gompertz distribution can be derived by realizing that its moment-generating function can be represented by the generalized integro-exponential function (Milgram, 1985). Unfortunately, despite its simple looking hazard function,

$$h(x) = ae^{bx} \quad a > 0,$$

the moments of the Gompertz distribution can only be formulated in terms of special functions. The nth moment of a Gompertz distributed random variable X is

$$E\left[X^n\right] = \frac{n!}{b^n} e^{\frac{a}{b}} E_1^{n-1}\left(\frac{a}{b}\right),$$

where $E_s^n(z) = \frac{1}{n!}\int_1^\infty (\ln x)^n x^{-s} e^{-zx}\, dz$ is the generalized integro-exponential function (Milgram, 1985). The advantage of using the generalized integro-exponential function is that it has

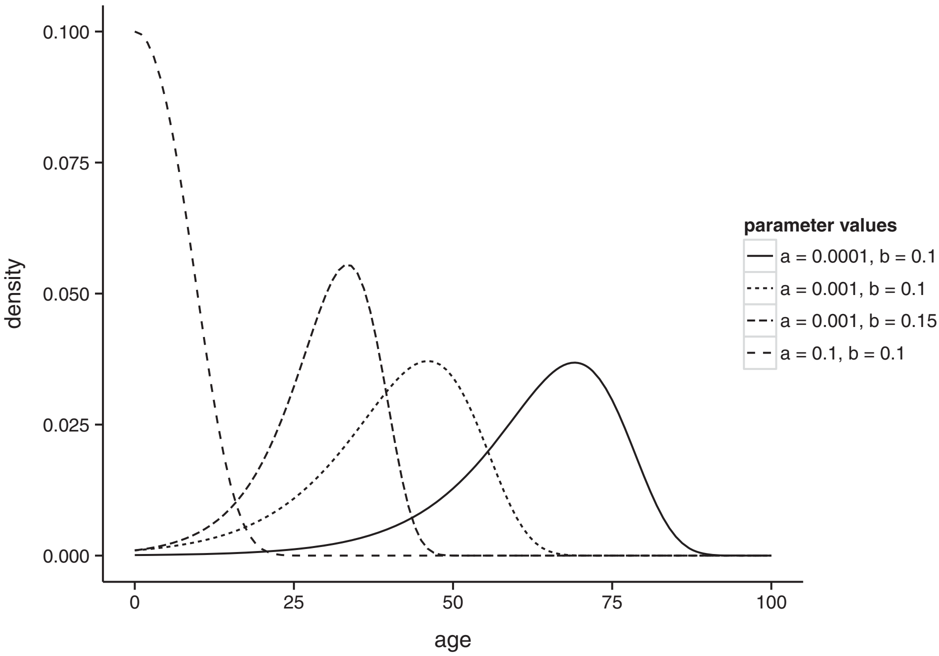

Figure 1. Shape of the Gompertz distribution. The Gompertz distribution for different combination of *a* and *b* parameters.

known power series expansion and also can be transformed to the succinct form of Meijer-G functions (Lenart, 2012).

$$E\left[X^n\right] = \frac{n!}{b^n} e^{\frac{a}{b}} G^{n+1,0}_{n,n+1}\left(\frac{a}{b}\,\middle|\, \begin{matrix} ;1,\ldots,1 \\ 0,\ldots,0; \end{matrix}\right),$$

where the Meijer G-function is a generalized hypergeometric function. It is defined by the contour integral

$$G^{m,n}_{p,q}\left[z\,\middle|\,\begin{matrix} a_1,\ldots,a_n;\,a_{n+1},\ldots,a_p \\ b_1,\ldots,b_m;\,b_{m+1},\ldots,b_q \end{matrix}\right] = \frac{1}{2\pi i}\int_C \frac{\prod_{j=1}^{m}\Gamma(b_j-s)\prod_{j=1}^{n}\Gamma\left(1-a_j+s\right)}{\prod_{j=m+1}^{q}\Gamma(1-b_j+s)\prod_{j=n+1}^{p}\Gamma\left(a_j-s\right)} z^s\,ds$$

along contour C (Erdélyi, 1953).

An interesting property of the Gompertz distribution is that the distribution can be truncated at any x and by rescaling the a parameter, the distribution will still yield a proper density function (Garg et al., 1970). Therefore, when studying, for example, the remaining life expectancy at $x > 0$, after rescaling the a parameter, the analyzed age will become the new 0.

1.1.1 *Relation to the generalized extreme value distribution*

The generalized extreme value distributions have the density function

$$f_{GEV}(x) = \frac{1}{\sigma}\left[1+\xi\left(\frac{x-m}{\sigma}\right)\right]^{-\left(\frac{1}{\xi}\right)-1} e^{-\left[1+\xi\left(\frac{x-m}{\sigma}\right)\right]^{-\frac{1}{\xi}}}, \qquad x\in\mathbb{R},$$

characterized by m location, σ scale and ξ shape parameters. For $\xi := 0$

$$f_{Gumbel}(x) = \frac{1}{\sigma} e^{-e^{-\frac{x-m}{\sigma}} - \frac{x-m}{\sigma}}, \qquad x \in \mathbb{R},$$

the generalized extreme value distribution degenerates into the Gumbel distribution. The Gumbel distribution is often used by hydrologists to calculate the probability of floods or extreme rainfall (e.g., Landwehr et al., 1979; Watterson and Dix, 2003). Formally, the Gompertz distribution is a special case of the Gumbel distribution for the minima, i.e., when $x := -x$ and truncated at $x = 0$ with Gompertz parameters substituted as $b = 1/\sigma$ and $a = b\exp(-bm)$:

$$f_{Gompertz}(x) = be^{b(x-m)+e^{-bm}-e^{b(x-m)}}, \qquad x \geq 0. \tag{2}$$

The Weibull distribution is another widespread distribution of the generalized extreme value family that is used in survival analysis (Lawless, 2011). The generalized extreme value distribution degenerates into the Weibull distribution for $\xi < 0$. The difference between the shape parameters govern the tail behavior of the distribution; the smaller the shape parameter, the thinner the tail is (Bali, 2003).

1.1.2 *Generalization of the Gompertz distribution*

A major drawback of the Gompertz distribution is that it fits only adult mortality sufficiently (Thatcher, 1999). After ages 80 or 90, the population level mortality starts to decelerate and the Gompertz hazard would overestimate the observed marginal hazard of the population. Vaupel et al. (1979) proposed to use a logistic, or gamma-Gompertz (GG), curve to provide a better fit:

$$h_{GG}(x) = \frac{ae^{bx}}{1 + \gamma \frac{a}{b}\left(e^{bx} - 1\right)} \qquad x \geq 0\,;\, a, \gamma > 0$$

to model the mortality deceleration above age 80. Here $\bar{\mu}(x)$ denotes the marginal hazard, or average hazard on the population level at age x. This improved model not only fits the data better but also provides a rationale for the slowing pace of mortality increase. They hypothesize that each individual is born with a level of frailty that increases or decreases their hazard of dying. Frailty can be interpreted as a random variable, if it is distributed according to the gamma distribution with same shape and scale parameters, then the average frailty of the population will be equal to 1 and the coefficient of variation of the gamma distribution, denoted by γ will be constant at all ages. As frailer individuals are more likely to die earlier than their more robust counterparts, the observed, marginal hazard levels off and mortality seems to decelerate.

2. Goodness-of-fit tests

2.1 *Correlation coefficient test*

Filliben (1975) introduced the probability plot coefficient test for normal distributions. The idea of the test is to compare the ordered observations with predicted order statistics of a theoretical distribution. Let $X_{[i]}$ denote the ith largest observed datum, $\widetilde{X}_{[i]}$ the order statistic median, $\bar{X}$ the average observation and $\widetilde{X}$ the population median, then the probability plot correlation coefficient is given by the Pearson correlation coefficient:

$$r = \frac{\sum_{i=1}^{n}\left(X_{[i]} - \bar{X}\right)\left(\widetilde{X}_{[i]} - \widetilde{X}\right)}{\sqrt{\sum\limits_{i=1}^{n}\left(X_{[i]} - \bar{X}\right)^2 \sum\limits_{i=1}^{n}\left(\widetilde{X}_{[i]} - \widetilde{X}\right)^2}}.$$

Filliben (1975) estimated the order statistic medians from the quantile function and later the same approach was used for the Gumbel and other extreme-value distributions (Vogel, 1986; Kinnison, 1989). These approaches relied on numerical approximations to the plotting positions between the order statistics and the order statistic medians or other measures of location[1] such as the plotting position of Gringorten (1963) which is unbiased only for the largest observation.

The correlation coefficient test can be improved by comparing the ordered observations with their expected values of a distribution. Let $X_{(i)}$ denote the ith smallest observation, $E[X_{(i)}]$ the expectation of it, and $E\left[X\right]$ the expected value of the theoretical population.

2.1.1 *Density and expected value of order statistics*

The density of $f_{(i)}(x)$ is (see, e.g., Harter, 1961)

$$f_{(i)}(x) = \frac{n!}{(i-1)!(n-i)!}F^{i-1}(x)\left(1 - F(x)\right)^{n-i} f(x)$$

and

$$E\left[X_{(i)}\right] = \int\limits_{-\infty}^{\infty} f_{(i)}(x)\, dx.$$

The density of $f_{(i)}(x)$ can be simplified by

$$X_{(i)} =_d F^{-1}\left(U_{(i)}\right), \tag{3}$$

where $U \sim U(0, 1)$ and F^{-1} is the quantile function of X. Because[2]2

$$f_{U_{(i)}}(x) = \frac{n!}{(i-1)!(n-i)!}x^{i-1}\left(1-x\right)^{n-i}, \qquad x \in [0, 1],$$

the expected value of $E[X_{(i)}]$ can be reformulated (Sen, 1959) as

$$E\left[X_{(i)}\right] = \frac{n!}{(i-1)!(n-i)!}\int_0^1 F^{-1}(x)x^{i-1}(1-x)^{n-i}\, dx.$$

2.1.2 *Correlation coefficient test for the Gompertz distribution*

The correlation coefficient test has the null hypothesis

$$H_0 \,:\, F(x) = G(x;\theta).$$

If $X \sim Gompertz(a, b)$, then

$$F^{-1}(x) = \frac{1}{b}\log\left(1 - \frac{b}{a}\log(1-x)\right), \qquad a, b > 0$$

[1] Plotting $X_{[i]}$ against $M_{[i]}$ yields an approximately linear plot.

[2] Note that the distribution function, $F_{U_{(i)}}(x)$ of the ith observation of a uniform distribution would be equal to the regularized incomplete beta function, $I_x(i, n-i+1)$ (Abramowitz and Stegun, 1965, 26.5).

and

$$E\left[X_{(i)}\right] = \frac{n!}{b(i-1)!(n-i)!} \int_0^1 \log\left(1 - \frac{b}{a}\log(1-x)\right) x^{i-1}(1-x)^{n-i}\, dx.$$

The expected value of the population is (Missov and Lenart, 2011)

$$E\left[X\right] = \frac{1}{b} e^{\frac{a}{b}} E_1\left(\frac{a}{b}\right),$$

where $E_n(z) = \int_1^\infty \exp(-zt)/t^n\, dt$ denotes the exponential integral (Abramowitz and Stegun, 1965, 5.1.4).

The estimated correlation coefficient is then

$$\hat{r}\left(\hat{\theta}\right) = \frac{\sum_{i=1}^n \left(X_{(i)} - \bar{X}\right)\left(E\left[X_{(i)}; \hat{\theta}\right] - E\left[X; \hat{\theta}\right]\right)}{\sqrt{\sum_{i=1}^n \left(X_{(i)} - \bar{X}\right)^2 \sum_{i=1}^n \left(E\left[X_{(i)}; \hat{\theta}\right] - E\left[X; \hat{\theta}\right]\right)^2}},$$

where $\hat{\theta}$ is the maximum likelihood estimate of $\theta = (a, b)$. The test statistic ranges from $[0, 1]$ and the null hypothesis is rejected if $\hat{r}$ is lower than a critical value estimated by Monte Carlo simulations (Table 1).

2.2 Anderson–Darling test

The Anderson and Darling (1952) test is based on the difference between the empirical and the theoretical distribution function $F(x)$ and $G(x)$,

$$W^2 = n \int_{-\infty}^{\infty} \left[F(x) - G(x)\right]^2 \psi(x)\, dG(x),$$

where $\psi(x)$ is a weight function. As Anderson and Darling (1952, p. 194) notes, for $\psi(x) := 1$ W^2 will be the same as the Cramér-von Mises test statistic

$$T = \frac{1}{12n} + \sum_{i=1}^n \left\{\frac{2i-1}{2n} - G\left[X_{(i)}\right]\right\}^2,$$

where $X_{(i)}$ is the ith smallest observation (Stephens, 1974). Other weight functions are also used to test the goodness-of-fit of extreme value distributions (e.g., Stephens, 1977), most notably $\psi(x) := \{G(x)\left[1 - G(x)\right]\}^{-1}$ that gives the Anderson–Darling test statistic (Shin et al., 2011)

$$\begin{aligned} A^2 &= n \int_{-\infty}^{\infty} \frac{\left[F(x) - G(x)\right]^2}{G(x)\left[1 - G(x)\right]}\, dG(x) \\ &= -n - \frac{1}{n} \sum_{i=1}^n (2i-1)\left\{\log G\left(X_{(i)}\right) + \log\left[1 - G\left(X_{(n-i+1)}\right)\right]\right\}. \end{aligned} \tag{4}$$

2.2.1 Extensions of the Anderson–Darling test

For testing the mortality of heterogeneous populations, the modified Anderson–Darling test statistic (Sinclair et al., 1990) is of interest. It attributes a different weight function for the upper and the lower tail

$$AU^2 = n \int_{-\infty}^{\infty} \frac{\left[F(x) - G(x)\right]^2}{1 - G(x)}\, dG(x)$$

Table 1. Correlation coefficient statistic. Empirical critical values of the correlation coefficient statistic.

		$a = 0.000001$				$a = 0.0001$				$a = 0.01$	
n	α	$b = 0.08$	$b = 0.10$	$b = 0.12$	$b = 0.14$	$b = 0.08$	$b = 0.10$	$b = 0.12$	$b = 0.14$	$b = 0.08$	$b = 0.10$
25	0.10	0.9576	0.9576	0.9576	0.9576	0.9606	0.9601	0.9599	0.9596	0.9759	0.9754
	0.05	0.9462	0.9462	0.9462	0.9461	0.9509	0.9503	0.9499	0.9495	0.9708	0.9702
	0.01	0.9160	0.9160	0.9160	0.9158	0.9280	0.9266	0.9259	0.9249	0.9584	0.9575
50	0.10	0.9730	0.9730	0.9730	0.9730	0.9764	0.9760	0.9757	0.9753	0.9877	0.9874
	0.05	0.9652	0.9651	0.9652	0.9651	0.9709	0.9703	0.9698	0.9693	0.9852	0.9849
	0.01	0.9422	0.9419	0.9417	0.9418	0.9579	0.9566	0.9557	0.9547	0.9794	0.9790
75	0.10	0.9796	0.9795	0.9795	0.9795	0.9830	0.9826	0.9823	0.9820	0.9917	0.9915
	0.05	0.9734	0.9734	0.9733	0.9733	0.9792	0.9786	0.9782	0.9778	0.9901	0.9899
	0.01	0.9548	0.9545	0.9544	0.9543	0.9705	0.9694	0.9685	0.9678	0.9864	0.9861
100	0.10	0.9833	0.9833	0.9833	0.9833	0.9867	0.9863	0.9860	0.9858	0.9937	0.9936
	0.05	0.9782	0.9782	0.9781	0.9781	0.9839	0.9833	0.9829	0.9825	0.9925	0.9924
	0.01	0.9627	0.9625	0.9622	0.9620	0.9773	0.9764	0.9757	0.9751	0.9898	0.9896
150	0.10	0.9875	0.9875	0.9875	0.9875	0.9907	0.9904	0.9902	0.9899	0.9958	0.9957
	0.05	0.9837	0.9836	0.9836	0.9836	0.9888	0.9884	0.9880	0.9877	0.9950	0.9949
	0.01	0.9721	0.9718	0.9716	0.9714	0.9843	0.9837	0.9832	0.9828	0.9932	0.9931
200	0.10	0.9896	0.9896	0.9896	0.9895	0.9928	0.9925	0.9923	0.9921	0.9968	0.9968
	0.05	0.9865	0.9865	0.9864	0.9864	0.9913	0.9910	0.9907	0.9905	0.9962	0.9962
	0.01	0.9774	0.9772	0.9769	0.9768	0.9879	0.9874	0.9870	0.9867	0.9949	0.9948
300	0.10	0.9904	0.9905	0.9905	0.9905	0.9947	0.9944	0.9942	0.9941	0.9977	0.9977
	0.05	0.9882	0.9882	0.9882	0.9882	0.9936	0.9934	0.9932	0.9930	0.9973	0.9973
	0.01	0.9819	0.9817	0.9816	0.9815	0.9913	0.9910	0.9907	0.9904	0.9965	0.9964

		$a = 0.01$		$a = 0.1$				$a = 0.2$			
n	α	$b = 0.12$	$b = 0.14$	$b = 0.08$	$b = 0.10$	$b = 0.12$	$b = 0.14$	$b = 0.08$	$b = 0.10$	$b = 0.12$	$b = 0.14$
25	0.10	0.9749	0.9745	0.9758	0.9761	0.9763	0.9765	0.9743	0.9746	0.9748	0.9750
	0.05	0.9696	0.9691	0.9702	0.9706	0.9709	0.9713	0.9682	0.9686	0.9689	0.9692
	0.01	0.9567	0.9560	0.9561	0.9570	0.9575	0.9581	0.9527	0.9533	0.9538	0.9543
50	0.10	0.9872	0.9869	0.9862	0.9866	0.9869	0.9872	0.9846	0.9850	0.9853	0.9857
	0.05	0.9846	0.9843	0.9830	0.9836	0.9841	0.9844	0.9809	0.9814	0.9819	0.9823
	0.01	0.9786	0.9781	0.9751	0.9762	0.9769	0.9776	0.9713	0.9722	0.9731	0.9739
75	0.10	0.9914	0.9912	0.9902	0.9906	0.9909	0.9911	0.9887	0.9891	0.9895	0.9898
	0.05	0.9897	0.9895	0.9880	0.9885	0.9889	0.9892	0.9860	0.9866	0.9871	0.9875
	0.01	0.9858	0.9855	0.9825	0.9834	0.9841	0.9847	0.9789	0.9799	0.9809	0.9816
100	0.10	0.9935	0.9934	0.9923	0.9927	0.9929	0.9931	0.9910	0.9914	0.9918	0.9921
	0.05	0.9923	0.9921	0.9907	0.9911	0.9915	0.9917	0.9889	0.9894	0.9899	0.9903
	0.01	0.9894	0.9892	0.9865	0.9873	0.9879	0.9883	0.9833	0.9843	0.9852	0.9858
150	0.10	0.9956	0.9956	0.9947	0.9949	0.9951	0.9953	0.9935	0.9939	0.9942	0.9944
	0.05	0.9948	0.9947	0.9935	0.9939	0.9941	0.9943	0.9920	0.9925	0.9929	0.9932
	0.01	0.9930	0.9928	0.9907	0.9913	0.9918	0.9921	0.9881	0.9890	0.9897	0.9902
200	0.10	0.9967	0.9966	0.9959	0.9961	0.9963	0.9964	0.9949	0.9952	0.9955	0.9957
	0.05	0.9961	0.9960	0.9950	0.9953	0.9955	0.9957	0.9938	0.9942	0.9945	0.9948
	0.01	0.9947	0.9946	0.9929	0.9934	0.9937	0.9940	0.9907	0.9915	0.9921	0.9925
300	0.10	0.9977	0.9976	0.9971	0.9973	0.9974	0.9975	0.9964	0.9967	0.9969	0.9970
	0.05	0.9973	0.9972	0.9965	0.9968	0.9969	0.9970	0.9956	0.9959	0.9962	0.9964
	0.01	0.9964	0.9963	0.9951	0.9955	0.9957	0.9959	0.9935	0.9941	0.9946	0.9949

$$= \frac{n}{2} - 2\sum_{i=1}^{n} G\left(X_{(i)}\right) - \sum_{i=1}^{n}\left(2 - \frac{2i-1}{n}\right)\log\left[1 - G\left(X_{(i)}\right)\right] \tag{5}$$

and

$$\begin{aligned} AL^2 &= n\int_{-\infty}^{\infty} \frac{[F(x) - G(x)]^2}{G(x)}\, dG(x) \\ &= -\frac{3n}{2} + 2\sum_{i=1}^{n} G\left(X_{(i)}\right) - \sum_{i=1}^{n} \frac{2i-1}{n} \log G\left(X_{(i)}\right), \end{aligned} \tag{6}$$

respectively. In a model where individuals have different levels of frailty (Vaupel et al., 1979) that acts multiplicatively on their baseline level of mortality, there would be more robust individuals (lower level of frailty) that would deviate in the upper tail from the homogeneous (all individuals having the same frailty) distribution.

2.2.2 *Anderson–Darling test for the Gompertz distribution*

As previously, the null hypothesis of the Anderson–Darling test is

$$H_0 \;:\; F(x) = G(x; \theta).$$

In case of the Gompertz distribution, $\theta = (a, b)$. By substituting

$$G(x; a, b) = 1 - e^{-\frac{a}{b}\left(e^{bx}-1\right)}$$

in either (4), (5), or (6), the Anderson–Darling test statistic is immediately given. Large values of the statistic reject the null hypothesis. The critical values are defined by Monte Carlo simulations (Table 2).

2.3 *Moments test for the Gompertz distribution*

An interesting, yet not very popular, goodness-of-fit test using moments was suggested by Li and Papadopoulos (2002). Suppose $X_1, \ldots, X_n$ are i.i.d. random variables characterized by a c.d.f. $F(x)$. We test a null hypothesis

$$H_0 : \text{F belongs to a parametric family } F_\theta,\ \theta \in \Theta$$

Suppose the k-th ($k \in \mathbb{N}$) moment $m_k = \int x^k dF_\theta(x)$ of F_θ exists and

$$g(m_1, \ldots, m_k) = 0 \qquad \forall\theta \in \Theta$$

for some function g. Then

$$\sqrt{n}\, g(\hat{m}_1, \ldots, \hat{m}_k) \rightarrow_d N(0, V(\theta))$$

$\hat{m}_i = \sum_{j=1}^{n} X_j^i/n$ denotes the sample moment of order i ($i = 1, \ldots, k$) and

$$V(\theta) = \nabla g(m_1, \ldots, m_k)^T\ \Sigma\ \nabla g(m_1, \ldots, m_k),$$

where $\Sigma = ||\sigma_{ij}||_{i,j=1}^k$ has elements $\sigma_{ij} = m_{i+j} - m_i m_j$ and $\nabla g(m_1, \ldots, m_k)$ denotes the gradient of g. We can choose $g(x, y, z) = z - 3xy + x^3$ and construct the following statistic:

$$T = \frac{\sqrt{n}\,(\hat{m}_3 - 3\hat{m}_1\hat{m}_2 + 2\hat{m}_1^3)}{\sqrt{V(\hat{a}, \hat{b})}} \underset{n\to\infty}{\longrightarrow} N(0, 1),$$

where $\hat{a}$ and $\hat{b}$ are the maximum likelihood estimates of the Gompertz parameters. For m_i, $i = 1, \ldots, 6$, we use the expressions calculated in Lenart (2012).

2.4 *Nested test against the truncated generalized extreme value distribution for the minima*

Let $f_{GEV}(x)$ be truncated at $x = 0$ (Elandt-Johnson, 1976), then

$$f_{tGEV}(x) = \frac{1}{\sigma}\left[1 + \xi\frac{(-x-m)}{\sigma}\right]^{-\frac{1}{\xi}-1} \exp\left\{\left(1 - \xi\frac{m}{\sigma}\right)^{-\frac{1}{\xi}} - \left[1 + \xi\frac{(-x-m)}{\sigma}\right]^{-\frac{1}{\xi}}\right\} \quad (7)$$

Table 2. Anderson–Darling statistic. Empirical critical values of the Anderson–Darling statistic.

		$a = 0.000001$				$a = 0.0001$				$a = 0.01$	
n	α	$b = 0.08$	$b = 0.10$	$b = 0.12$	$b = 0.14$	$b = 0.08$	$b = 0.10$	$b = 0.12$	$b = 0.14$	$b = 0.08$	$b = 0.10$
25	0.10	0.6262	0.6263	0.6261	0.6262	0.6236	0.6242	0.6239	0.6245	0.6287	0.6264
	0.05	0.7428	0.7431	0.7428	0.7428	0.7402	0.7407	0.7405	0.7410	0.7511	0.7477
	0.01	1.0098	1.0097	1.0096	1.0100	1.0079	1.0069	1.0068	1.0076	1.0393	1.0321
50	0.10	0.6299	0.6300	0.6299	0.6301	0.6269	0.6274	0.6273	0.6279	0.6322	0.6294
	0.05	0.7491	0.7491	0.7487	0.7491	0.7448	0.7461	0.7459	0.7465	0.7557	0.7518
	0.01	1.0213	1.0219	1.0213	1.0220	1.0170	1.0182	1.0182	1.0189	1.0444	1.0380
75	0.10	0.6315	0.6315	0.6311	0.6314	0.6282	0.6285	0.6289	0.6290	0.6335	0.6301
	0.05	0.7513	0.7513	0.7505	0.7509	0.7475	0.7480	0.7484	0.7486	0.7573	0.7532
	0.01	1.0258	1.0267	1.0263	1.0258	1.0231	1.0226	1.0232	1.0228	1.0479	1.0404
100	0.10	0.6323	0.6322	0.6319	0.6319	0.6287	0.6294	0.6291	0.6297	0.6339	0.6306
	0.05	0.7522	0.7526	0.7523	0.7525	0.7482	0.7487	0.7492	0.7494	0.7578	0.7531
	0.01	1.0286	1.0288	1.0298	1.0281	1.0237	1.0234	1.0234	1.0247	1.0480	1.0390
150	0.10	0.6328	0.6328	0.6325	0.6330	0.6290	0.6296	0.6300	0.6302	0.6341	0.6311
	0.05	0.7529	0.7529	0.7530	0.7532	0.7493	0.7493	0.7504	0.7503	0.7577	0.7540
	0.01	1.0305	1.0311	1.0296	1.0301	1.0248	1.0251	1.0276	1.0262	1.0485	1.0412
200	0.10	0.6330	0.6331	0.6334	0.6332	0.6293	0.6302	0.6302	0.6304	0.6347	0.6315
	0.05	0.7536	0.7539	0.7539	0.7538	0.7493	0.7501	0.7502	0.7508	0.7587	0.7541
	0.01	1.0309	1.0298	1.0318	1.0311	1.0261	1.0266	1.0268	1.0273	1.0492	1.0397
300	0.10	0.6334	0.6338	0.6333	0.6332	0.6301	0.6301	0.6306	0.6310	0.6351	0.6317
	0.05	0.7539	0.7547	0.7543	0.7541	0.7502	0.7502	0.7507	0.7515	0.7587	0.7541
	0.01	1.0324	1.0321	1.0321	1.0316	1.0260	1.0274	1.0287	1.0285	1.0493	1.0413
500	0.10	0.6333	0.6335	0.6333	0.6335	0.6298	0.6308	0.6310	0.6315	0.6356	0.6324
	0.05	0.7543	0.7544	0.7543	0.7544	0.7502	0.7508	0.7515	0.7520	0.7598	0.7551
	0.01	1.0325	1.0331	1.0326	1.0322	1.0283	1.0284	1.0290	1.0299	1.0493	1.0412
1000	0.10	0.6337	0.6338	0.6342	0.6340	0.6297	0.6310	0.6310	0.6313	0.6353	0.6325
	0.05	0.7546	0.7547	0.7554	0.7550	0.7502	0.7519	0.7516	0.7519	0.7592	0.7557
	0.01	1.0333	1.0340	1.0329	1.0347	1.0268	1.0302	1.0299	1.0306	1.0492	1.0419

		$a = 0.01$		$a = 0.1$				$a = 0.2$			
n	α	$b = 0.12$	$b = 0.14$	$b = 0.08$	$b = 0.10$	$b = 0.12$	$b = 0.14$	$b = 0.08$	$b = 0.10$	$b = 0.12$	$b = 0.14$
25	0.10	0.6246	0.6238	0.6802	0.6752	0.6725	0.6684	0.7055	0.7020	0.6976	0.6952
	0.05	0.7450	0.7437	0.8243	0.8170	0.8128	0.8079	0.8595	0.8543	0.8483	0.8447
	0.01	1.0268	1.0235	1.1755	1.1635	1.1560	1.1453	1.2423	1.2321	1.2225	1.2132
50	0.10	0.6278	0.6262	0.6891	0.6828	0.6777	0.6735	0.7124	0.7070	0.7022	0.6978
	0.05	0.7492	0.7470	0.8344	0.8258	0.8189	0.8130	0.8657	0.8591	0.8523	0.8465
	0.01	1.0328	1.0289	1.1855	1.1683	1.1577	1.1463	1.2394	1.2279	1.2142	1.2045
75	0.10	0.6282	0.6268	0.6943	0.6867	0.6810	0.6764	0.7166	0.7110	0.7048	0.7002
	0.05	0.7502	0.7476	0.8403	0.8303	0.8226	0.8160	0.8705	0.8632	0.8549	0.8486
	0.01	1.0346	1.0295	1.1903	1.1745	1.1601	1.1499	1.2430	1.2308	1.2167	1.2059
100	0.10	0.6288	0.6270	0.6959	0.6892	0.6826	0.6774	0.7195	0.7122	0.7066	0.7021
	0.05	0.7503	0.7481	0.8428	0.8335	0.8240	0.8175	0.8735	0.8640	0.8572	0.8507
	0.01	1.0350	1.0299	1.1921	1.1778	1.1629	1.1511	1.2446	1.2297	1.2185	1.2061
150	0.10	0.6294	0.6279	0.6987	0.6907	0.6848	0.6795	0.7225	0.7156	0.7100	0.7040
	0.05	0.7510	0.7493	0.8461	0.8351	0.8270	0.8193	0.8780	0.8681	0.8607	0.8529
	0.01	1.0362	1.0318	1.1964	1.1776	1.1661	1.1533	1.2500	1.2344	1.2225	1.2088
200	0.10	0.6294	0.6283	0.7001	0.6921	0.6857	0.6792	0.7239	0.7168	0.7106	0.7052
	0.05	0.7516	0.7498	0.8473	0.8368	0.8277	0.8194	0.8794	0.8699	0.8615	0.8542
	0.01	1.0363	1.0332	1.1981	1.1811	1.1665	1.1541	1.2519	1.2357	1.2231	1.2113
300	0.10	0.6296	0.6282	0.7016	0.6935	0.6859	0.6808	0.7261	0.7185	0.7116	0.7063
	0.05	0.7517	0.7497	0.8492	0.8382	0.8285	0.8213	0.8820	0.8713	0.8620	0.8557
	0.01	1.0361	1.0319	1.2002	1.1819	1.1683	1.1542	1.2556	1.2373	1.2218	1.2124
500	0.10	0.6302	0.6286	0.7026	0.6942	0.6871	0.6815	0.7280	0.7209	0.7136	0.7078
	0.05	0.7521	0.7498	0.8504	0.8384	0.8295	0.8221	0.8845	0.8742	0.8650	0.8575
	0.01	1.0372	1.0333	1.2033	1.1833	1.1686	1.1552	1.2592	1.2397	1.2267	1.2137
1000	0.10	0.6300	0.6284	0.7036	0.6947	0.6887	0.6826	0.7293	0.7213	0.7145	0.7091
	0.05	0.7517	0.7502	0.8518	0.8395	0.8314	0.8235	0.8857	0.8751	0.8655	0.8588
	0.01	1.0358	1.0338	1.2037	1.1825	1.1700	1.1571	1.2596	1.2415	1.2268	1.2159

Table 3. Alternative distributions. Density and support of alternative distributions.

Distribution	Density	Support
Weibull(x;a,b)	$\frac{a}{b}\left(\frac{x}{b}\right)^{a-1} e^{-\left(\frac{x}{b}\right)^a}$	$[0, \infty)$
Log-normal(x;μ,σ)	$\frac{1}{x\sigma\sqrt{2\pi}} e^{-\frac{(\log(x)-\mu)^2}{2\sigma^2}}$	$(0, \infty)$
Normal(x;μ,σ)	$\frac{1}{\sigma\sqrt{2\pi}} e^{-\frac{(x-\mu)^2}{2\sigma^2}}$	$(-\infty, \infty)$
Truncated Normal(x;μ,σ,α,β)	$\phi(x;\mu,\sigma)/[\Phi(\beta,\mu,\sigma)-\Phi(\alpha,\mu,\sigma)]^*$	$[\alpha, \beta]$
Logistic(x;μ,σ)	$\frac{1}{\sigma} e^{\frac{x-\mu}{\sigma}}\left(1+e^{\frac{x-\mu}{\sigma}}\right)^{-2}$	$(-\infty, \infty)$
Log-logistic(x;a,b)	$\frac{b}{a}\left(\frac{x}{a}\right)^{b-1} / \left[1+\left(\frac{x}{a}\right)^b\right]^2$	$[0, \infty)$
Inverse Gaussian(x;μ,λ)	$\left(\frac{\lambda}{2\pi x^3}\right)^{\frac{1}{2}} e^{\frac{-\lambda(x-\mu)^2}{2\mu^2 x}}$	$(0, \infty)$
Gamma(x;k,σ)	$\frac{1}{\sigma^k\Gamma(k)} x^{k-1} e^{-\frac{x}{\sigma}}$	$(0, \infty)$
Gamma–Gompertz(x;a,b,γ)	$\frac{ae^{bx}}{1+\gamma\frac{a}{b}(e^{bx}-1)}\left[1+\gamma\frac{a}{b}\left(e^{bx}-1\right)\right]^{-\frac{1}{\gamma}}$	$[0, \infty)$

*$\phi(\cdot)$ and $\Phi(\cdot)$ denote the normal density and distribution functions, respectively.

is the density function of the truncated generalized extreme value distribution for the minima with support on $[0, \infty)$, where m is the location, ξ is the shape, and σ is the scale parameter.

To test whether the Gompertz distribution fits the data as well as the truncated generalized extreme value distribution for the minima,

$$H_0 : \xi = 0$$

a likelihood ratio test is employed

$$-2\log\frac{L\left(g(x;\hat{a},\hat{b})\right)}{L\left(f_{tGEV}(x;\hat{a},\hat{b},\hat{\xi})\right)} \sim \chi^2(1),$$

where $L(\cdot)$ denotes the likelihood function and $g(\cdot)$ the Gompertz distribution. The likelihood ratio is evaluated at the maximum likelihood estimates of the two log-likelihood functions and by Wilks (1938) the limiting distribution of the likelihood ratio test statistic is the χ^2 distribution with degrees of freedom equal to the number of constraints under the null hypothesis.

3. Power of the tests

To compare the tests, $n = 50$ and $n = 200$ samples were simulated from alternative distributions repeated 50,000 times each. For the density and support of the alternative distributions please see Table 3. These alternative distributions were Weibull, log-normal, normal, logistic, gamma, and GG distributions. As the main application area of the Gompertz distribution is the analysis of life times, two sets of parameter values of the alternatives were each chosen as likely parameters describing current human longevity distributions with modal age at death and life expectancy (i.e., expected value) either about 80–85 years (Canudas-Romo, 2000) or remaining life expectancy of about 5 years and Gompertz b parameter 0.1–0.13 (Canudas-Romo, 2000; Barbi, 2003). The former case might correspond to a 0 starting age of observation and it is termed negatively skewed Gompertz in Tables 4 and 5. The latter case might rather describe situations when the youngest observed individual is 85 years old and the corresponding power comparisons can be found under positively skewed Gompertz in Tables 4 and 5.

Table 4. Small sample power comparisons. Power of the goodness-of-fit statistics against alternative distributions with $n = 50, \alpha = 0.05$.

Alternatives for a negatively skewed Gompertz distribution	r	AD	M	LR
Weibull(10,80)	0.0497	0.0934	0.0974	0.0982
Log-normal(4.4,0.01)	0.4793	0.7855	0.3124	0.3965
Normal(80,10)	0.5147	0.5524	0.3791	0.4014
Logistic(80,5)	0.4895	0.6870	0.4101	0.4534
Log-logistic(81,15)	0.7697	0.9999	0.6711	0.6828
Inverse Gaussian(81,4554)	0.8264	0.8392	0.6721	0.6983
Gamma(71,1.1)	0.6863	0.7390	0.2710	0.5747
Gamma-Gompertz(0.001,0.1,0.2)	0.0892	0.0836	0.0869	0.0955
Gompertz(0.0002,0.12)	0.2632	0.0522	0.0571	0.0516
Alternatives for a positively skewed Gompetz distribution	r	AD	M	LR
Weibull(1.5,6)	0.2591	0.2274	0.2481	0.2733
Truncated Normal(3,6,0,∞)	0.0943	0.0621	0.0661	0.0558
Inverse Gaussian(6,1.6)	0.9590	0.9734	0.8645	0.0150
Log-logistic(5,1.8)	0.9958	0.7845	0.9176	1.0000
Gamma(1.5,0.25)	0.3102	0.2124	0.1774	0.1145
Gamma-Gompertz(0.1,0.1,0.2)	0.1048	0.0604	0.0579	0.0002
Gompertz(0.1,0.1)	0.0512	0.0531	0.0506	0.0305

The Weibull distribution is an asymmetric distribution often used in survival analysis and reliability engineering. The mode of Weibull(10,80) is 79.2 with fitted Gompertz $b \approx 0.125$. The expected value of Weibull(1.5,6) that corresponds to the positively skewed Gompertz case is 5.4

The log-normal distribution is also asymmetrical and used as a statistical model for life times (Lawless, 2011). The modal age at death in a log-normal(4.4,0.01) life time distribution would be about 81.4 and estimate of Gompertz $b \approx 0.11$. As its density function is not likely to characterize an observed density of a positively skewed Gompertz distribution, it was dropped from the list of alternatives to test the power of positively skewed Gompertz distributions.

Adult life times were often assumed to follow a normal distribution (Véron and Rohrbasser (2003) citing Wilhelm Lexis) with standard deviation 9.3 (Ediev (2012) citing Wilhelm Lexis)

Table 5. Larger sample power comparisons. Power of the goodness-of-fit statistics against alternative distributions with $n = 200, \alpha = 0.05$.

Alternatives for a negatively skewed Gompertz distribution	r	AD	M	LR
Weibull(10,80)	0.1460	0.2843	0.3542	0.5149
Log-normal(4.4,0.01)	1.0000	1.0000	0.6813	0.9491
Normal(80,10)	0.9863	0.9950	0.5631	0.9687
Logistic(80,5)	0.9634	0.9983	0.4773	0.9676
Log-logistic(81,15)	1.000	1.000	0.9396	0.9879
Inverse Gaussian(81,4554)	0.9679	1.0000	0.8013	0.9873
Gamma(71,1.1)	0.9999	0.9089	0.8129	0.9632
Gamma-Gompertz(0.001,0.1,0.2)	0.0287	0.3009	0.3939	0.2692
Gompertz(0.0002,0.12)	0.0484	0.0518	0.0573	0.0590
Alternatives for a positively skewed Gompetz distribution	r	AD	M	LR
Weibull(1.5,6)	0.6825	0.8286	0.7451	0.8737
Truncated Normal(3,6,0,∞)	0.1785	0.1313	0.1243	0.1622
Inverse Gaussian(6,1.6)	1.0000	1.0000	0.9182	0.1801
Log-logistic(5,1.8)	0.9995	0.9999	0.9648	1.0000
Gamma(1.5,0.25)	0.6999	0.7851	0.6465	0.6682
Gamma-Gompertz(0.1,0.1,0.2)	0.1722	0.0802	0.0731	0.0688
Gompertz(0.1,0.1)	0.0502	0.0512	0.0521	0.0567

and modal age at death 80 for modern populations. However, in the case a positively skewed Gompertz distribution as the support of the normal distribution is on $(-\infty, \infty)$, a significant portion of a likely alternative normal distribution's probability density would be on the negative axis. Therefore, instead of the normal distribution, a truncated normal distribution from below at 0 was used as an alternative.

The logistic distribution is often cited as the observed shape of the hazard function in many biological studies (Wilson, 1994) and logistic (80,5) yield a similar but less dispersed distribution of life times as the normal (80,10). However, similarly to the normal distribution, it has support on the whole real axis and cannot be used as an alternative for the positively skewed Gompertz case.

The log-logistic and inverse Gaussian distributions are also sometimes used as survival distributions (Folks and Chhikara, 1978; Bennett, 1983). The log-logistic(81,15), log-logistic(5,1.8), inverse Gaussian(81,4554), and inverse Gaussian(6,1.6) distributions have a modal value of 80.3, 2.5, 81, and 5, respectively.

The gamma distribution has a flexible shape and is also used as a life time distribution (Lawless, 2011) with gamma (71,1.1) giving modal longevity of 77 and fitted Gompertz b parameter ≈ 0.12. The modal value of gamma (1.5,0.25) is 0.125. The GG distribution (Vaupel et al., 1979) is a generalized form of the Gompertz distribution with a logistic shape of the hazard. A GG (0.001,0.1,0.2) correspond to the distribution of remaining lifespan of modern populations at about age 70 (Missov, 2013). The GG (0.1,0.1,0.2) relate to the remaining life time of current populations at about age 85. Note that the normal and the logistic distributions are the only symmetric distributions among the alternatives for the power comparison.

The most powerful test was the Anderson–Darling test for all except the Weibull and the GG distributions. Not surprisingly, the likelihood ratio test was the best to identify the differences between the Gompertz and the Weibull distribution and was also effective against the GG distribution. The modified Anderson–Darling test, with emphasis on the upper tail of the distribution could distinguish between Gompertz and GG distributions 12% of the samples of size 50.

The rejection rate of the tests increases for larger samples with the exception of the test for the sample mean. It seems that the most powerful tests for the Gompertz distribution are the Anderson–Darling and the correlation coefficient tests, especially if they tests against a less related distribution (log-normal, normal, logistic, or gamma). If the test is against a related distribution such as Weibull or GG, the efficiency of all tests drop. Against the Weibull distribution, the likelihood ratio against the generalized extreme value distribution works the best, its efficiency is lower for the GG model as the test is not explicitly against it. When the alternative is a positively skewed Weibull distribution, all of the tests perform better. It is more difficult to evaluate the power of the likelihood ratio test against non extreme value distributions. It has a relatively high rejection rate against all of the other distribution but it is not an appropriate test against them as they are not members of the family of extreme value distributions. It is especially apparent against the test of inverse Gaussian(6,1.6). The moments test performs best in the Weibull and GG cases, yet has the weakest power in all other settings.

4. Application: Goodness-of-fit to laboratory rat data

The goodness-of-fit tests defined above can be readily used to check if empirical data is Gompertz distributed. As an example, individual life span data of rats will be used. The analyzed data was collected by Vladimir N. Anisimov at the N.N. Petrov Research Institute of Oncology, St. Petersburg, Russia, to test carcinogenicity and it is now published in

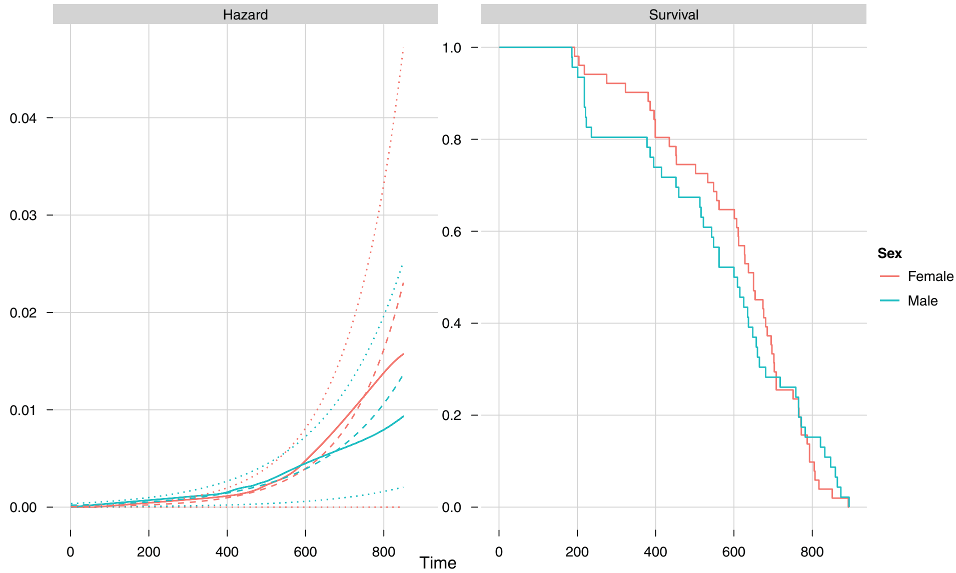

Figure 2. Hazard and survival of the rat data. On the left panel, the solid line corresponds to the non parametric hazard estimate, the dashed line to the Gompertz fit, and the dotted lines are the 95% confidence intervals of the fitted Gompertz hazard.

Table 6. Rat survival. Descriptive statistics of life spans of 51 female and 46 male rats (days).

Sex	n	Min	q_1	$\tilde{x}$	$\bar{x}$	q_3	Max	s	IQR
Female	51	192.5	477.0	649.5	603.2	729.0	891.5	177.9	252
Male	46	185.5	399.5	604.0	559.1	747.5	893.5	219.4	348

the Biodemographic Database (BDB). Here we will use only the rats in the control group, $n = 51$ females and $n = 46$ males. The data is fully observed and the number of survivors was recorded every day. Please see Fig. 2 for the estimated hazard and the Kaplan–Meier survival function and Table 6 for descriptive statistics of the dataset. The hazard estimation was carried out by the same varying kernel width estimation procedure as mentioned earlier. The Gompertz fit to the data show very wide confidence intervals which were estimated by the delta method.

The goodness-of-fit statistics in general do not reject the null hypothesis that both the distribution of death of both the male and the female rats is Gompertz (Table 7). While the maximum likelihood estimate of a of the male rats is higher than $\hat{a}$ of the female rats, the estimated daily rate of aging parameter, $\hat{b}$ is lower, leading to a cross-over of mortality later in life (Fig. 2). This result is corroborated by the non parametric estimates. However, because of the low sample size, the confidence bands are very wide. In spite of that, by looking at the goodness-of-fit statistics and their respective critical values in the Appendix, it can be seen

Table 7. Goodness-of-fit of the rat data. Calculated Gompertz goodness-of-fit test statistics to the dataset of 51 female and 46 male rats (in parentheses the associated p-values).

Sex	$\hat{a}$	$\hat{b}$	$\bar{\mu}_{\alpha=0.01}$	r	AD	M	LR
Female	5.7×10^{-5}	0.007	−0.0014–0.0029	0.991	0.384	−1.149	0.895
Male	1.9×10^{-3}	0.005	−0.00012–0.0034	0.983	0.55	−0.833	2.165

that the null is not rejected either by the Anderson–Darling ($0.384 < 0.63$ and $0.55 < 0.62$) and the correlation coefficient ($0.991 > 0.973$ and $0.983 > 0.976$) test statistics at $\alpha = 0.1$. The likelihood ratio test also confirms that the Gompertz distribution fits the data as well as the generalized extreme value distribution (its shape parameter equals to 0) at $\alpha = 0.1$ for both females ($0.895 < 2.71$) and males ($2.165 < 2.71$). The moments test similarly does not reject the null with $M = -1.149$ ($p = 0.251$) and $M = -0.833$ ($p = 0.405$) for females and males, respectively.

5. Discussion

The comparison of the power of the tests show that the Anderson–Darling statistic is the most powerful in rejecting the null that the empirical distribution comes from the Gompertz distribution when it was simulated from an alternative distribution. The Anderson–Darling statistic implemented by its computing formula is also the simplest and the quickest to run, and an important advantage of it is that for low values of a, the distribution of the statistic is independent from the Gompertz a and b parameters.

The correlation coefficient test also efficiently refutes other alternative distributions, however, when the alternative distribution is closely related to the Gompertz, such as in the case of Weibull and GG distributions, the power of the correlation coefficient test drops. As Legates and McCabe (1999) noted, the tests based on correlation are overly sensitive to outliers and insensitive to proportional differences between the expected and the observed values.

Juxtaposed with the results for the Gumbel distribution (Pérez-Rodríguez et al., 2009), the Kullback–Leibler test, not shown here, performs unexpectedly poorly relative to the other tests. The main disadvantage of the Kullback–Leibler test lies in the estimation of the sample entropy. The critical values obtained by the numerical procedure of Song (2002) vary substantially from dataset to dataset with similar sample sizes.

The likelihood ratio test is a powerful test when the alternative distribution is from the generalized extreme value family. A positive externality of the test is that the shape parameter of the generalized extreme value distribution, ξ has to be estimated during the testing procedure. If $\xi < 0$ and the likelihood ratio at the chosen significance level rejects the null hypothesis that $\xi = 0$, than the empirical distribution can be better fitted by a Weibull distribution than by a Gompertz. If $\xi > 0$, the empirical distribution is more likely to be Fréchet-type than Gompertz (Jenkinson, 1955).

Acknowledgments

The authors wish to thank Ulrich Halekoh and the anonymous reviewers for helpful comments that led to a significant improvement of the paper.

Appendix. Empirical critical values

The Gompertz distribution is a truncated Gumbel distribution for the minima. The Gumbel distribution is a member of the location-scale family of distributions, therefore its test statistics are independent of the location or scale parameters and simple Monte Carlo methods yield unbiased empirical critical values (Pérez-Rodríguez et al., 2009). However, as the Gompertz distribution is truncated from below at 0, its parameters are negatively correlated (Strehler and Mildvan, 1960; Lestienne, 1988) and it ceases to come from the location-scale family. Simply sampling from the distribution function would give biased critical values for the test statistics. In this case, the distribution of the test statistic should be simulated "after replacing the nuisance parameters by a consistent point estimate" (Dufour, 2006, 446) such as the maximum likelihood estimate.

Therefore, empirical critical values were calculated by parametric bootstrapping (Hall, 1992) where $N^* = 7,000$ samples were drawn from the Gompertz distribution for each combination of $a = \{0.000001, 0.0001, 0.01, 0.1, 0.2\}$, $b = \{0.08, 0.1, 0.12, 0.14\}$, and sample size $n = \{50, 75, 100, 150, 200, 300, 500, 1000\}$. Following a maximum likelihood estimation to each sample, $NB^* = 1,000$ samples were simulated from $Gompertz(\hat{a}, \hat{b})$ with the fitted parameters $\hat{a}$ and $\hat{b}$ and their respective correlation coefficient and Anderson–Darling test statistics were calculated.Finally, the empirical critical values of the test statistics were calculated as the means of the $(1-\alpha)$-quantiles of the test statistics. Algorithm 1 shows the structure of the simulations in pseudocode.

The empirical critical values show that the distribution of the test statistics are independent of parameter b when parameter a is relatively small compared to it as in this case the Gompertz distribution behaves as a Gumbel distribution for the minima. However, when parameter a becomes large relative to parameter b, the distribution of the test statistics depend on both parameters.

Algorithm 1 Calculation of empirical critical values by parametric bootstrapping

Require: $n > 0, a > 0, b \in \mathbb{R}, 0 < \alpha < 1, N^* > 0, NB^* > 0$
  Define vector of n, a, b, α, $N = 0$ and $NB = 0$
  for each n **do**
    for each a **do**
      for each b **do**
        repeat
          simulate Gompertz(a,b) of size n
          Fit Gompertz(a,b) by ML and obtain MLEs $\hat{a}$ and $\hat{b}$
          repeat
            simulate $Gompertz(\hat{a}, \hat{b})$ of size n
            calculate r and AD
            $NB \leftarrow NB + 1$
          until $NB = NB^*$
          $c_{r,n,a,b,\alpha,N} \leftarrow (1-\alpha)$-quantile of $r_{n,a,b}$
          $c_{AD,n,a,b,\alpha,N} \leftarrow (1-\alpha)$-quantile of $AD_{n,a,b}$
          $N \leftarrow N + 1$
        until $N = N^*$
        for each α **do**
          $c_{r,n,a,b,\alpha} \leftarrow \frac{1}{N}\sum_{i=1}^{N} c_{r,n,a,b,\alpha,N}$
          $c_{AD,n,a,b,\alpha} \leftarrow \frac{1}{N}\sum_{i=1}^{N} c_{AD,n,a,b,\alpha,N}$
        end for
      end for
    end for
  end for

A.1 *Critical values of the correlation coefficient statistic*

For sample sizes over 300, the critical values of the correlation coefficient statistic was omitted as the numerical computation of the statistic is not entirely reliable as it requires to calculate high values of factorials. In practice, the $n!/b(i-1)!(n-i)!$ term can be more efficiently calculated by $1/\beta(i, (n-i+1))$ as the beta function can be counted until higher values than the factorials separately. For even larger samples, the samples can be drawn from the quantile function by (3) by noting that the rank percentiles (rank of the observation divided by sample size $+1$) are also bounded by 0 and 1 (see, e.g., Kinnison, 1989).

A.2 *Critical values of the Anderson–Darling statistic*

Please note that the Anderson–Darling statistics are stable over all low values of $\hat{a}$ and increasing by $\hat{a}$. The critical values also increase slightly as the sample size increases. Similar trend was found by Shin et al. (2011) for the modified Anderson–Darling test.

References

Abramowitz, M., Stegun, I. (1965). *Handbook of Mathematical Functions*. Washington, DC: US Government Printing Office.
Anderson, T., Darling, D. (1952). Asymptotic theory of certain "goodness of fit" criteria based on stochastic processes. *Ann. Math. Stat.* 23:193–212.
Bali, T. (2003). The generalized extreme value distribution. *Econ. Lett.* 79:423–427.
Barbi, E. (2003). Trajectories of extreme survival in heterogeneous populations (English edition). *Population* 58:43–66.
Benjamin, B., Haycocks, H., Pollard, J. (1980). *The Analysis of Mortality and Other Actuarial Statistics*. London: Heinemann.
Bennett, S. (1983). Log-logistic regression models for survival data. *Appl. Stat.* (2):165–171.
Canudas-Romo, V. (2000). The modal age at death and the shifting mortality hypothesis. *Demogr. Res.* 19:1179–1204.
Doblhammer, G. (2000). Reproductive history and mortality later in life: a comparative study of england and wales and austria. *Popul. Stud.* 54:169–176.
Dufour, J.-M. (2006). Monte carlo tests with nuisance parameters: a general approach to finite-sample inference and nonstandard asymptotics. *J. Economet.* 133:443–477.
Ediev, D.M. (2012). A note on the compression of mortality. In: *Annual Meeting of the Population Association of America 2012, San Francisco*.
Elandt-Johnson, R. (1976). Conditional failure time distributions under competing risk theory with dependent failure times and proportional hazard rates. *Scand. Actuarial J.* 1976:37–51.
Erdélyi, A. (1953). *Higher Transcendental Functions* (Vol. 1). New York: McGraw-Hill.
Filliben, J. (1975). The probability plot correlation coefficient test for normality. *Technometrics* 17:111–117.
Finch, C., Pike, M. (1996). Maximum life span predictions from the Gompertz mortality model. *J. Gerontol. Ser. A: Biol. Sci. Med. Sci.* 51:B183.
Finch, C., Pike, M., Witten, M. (1990). Slow mortality rate accelerations during aging in some animals approximate that of humans. *Science (New York, NY)* 249:902.
Folks, J., Chhikara, R. (1978). The inverse Gaussian distribution and its statistical application – a review. *J. R. Stat. Soc. Ser. B (Methodol.)* (3):263–289.
Garg, M., Rao, B., Redmond, C. (1970). Maximum-likelihood estimation of the parameters of the Gompertz survival function. *J. R. Stat. Soc. Ser. C (Appl. Stat.)* 19:152–159.
Gringorten, I. (1963). A plotting rule for extreme probability paper. *J. Geophys. Res.* 68:813–814.
Hall, P. (1992). *The Bootstrap and Edgeworth Expansion*. New York: Springer.
Harter, H. (1961). Expected values of normal order statistics. *Biometrika* 48:151–165.
Hirsch, H., Peretz, B. (1984). Survival and aging of a small laboratory population of a marine mollusc, *Aplysia californica*. *Mech. Ageing Dev.* 27:43–62.

Honda, S., Matsuo, M. (1992). Lifespan shortening of the nematode *Caenorhabditis elegans* under higher concentrations of oxygen. *Mech. Ageing Dev.* 63:235–246.
Hosking, J. (1984). Testing whether the shape parameter is zero in the generalized extreme-value distribution. *Biometrika* 71:367–374.
Jenkinson, A. (1955). The frequency distribution of the annual maximum (or minimum) values of meteorological elements. *Quart. J. R. Meteorol. Soc.* 81:158–171.
Kinnison, R. (1989). Correlation coefficient goodness-of-fit test for the extreme-value distribution. *Am. Statist.* 43:98–100.
Landwehr, J., Matalas, N., Wallis, J. (1979). Probability weighted moments compared with some traditional techniques in estimating Gumbel parameters and quantiles. *Water Resources Res.* 15:1055–1064.
Lawless, J. F. (2011). *Statistical Models and Methods for Lifetime Data* (2nd ed.). Hoboken, NJ: Wiley-Interscience.
Legates, D., McCabe, Jr, G. (1999). Evaluating the use of "goodness-of-fit" measures in hydrologic and hydroclimatic model validation. *Water Resources Res.* 35:233–241.
Lenart, A. (2012). The moments of the Gompertz distribution and maximum likelihood estimation of its parameters. *Scand. Actuarial J.* doi:10.1080/03461238.2012.687697.
Lestienne, R. (1988). On the thermodynamical and biological interpretation of the Gompertzian mortality rate distribution. *Mech. Ageing Dev.* 42:197–214.
Li, G., Papadopoulos, A. (2002). A note on goodness of fit tests using moments. *Statistica* 1:71–86.
Milgram, M. (1985). The generalized integro-exponential function. *Math. Comput.* 44:443–458.
Missov, T., Lenart, A. (2011). Linking period and cohort life-expectancy linear increases in Gompertz proportional hazards models. *Demogr. Res.* 24:455–468.
Missov, T. I. (2013). Gamma-Gompertz life expectancy at birth. *Demogr. Res.* 28:259–270.
Pérez-Rodríguez, P., Vaquera-Huerta, H., Villaseñor-Alva, J. (2009). A goodness-of-fit test for the gumbel distribution based on Kullback–Leibler information. *Commun. Stat. Theory Methods* 38:842–855.
Perozek, M. (2008). Using subjective expectations to forecast longevity: do survey respondents know something we don't know? *Demography* 45:95–113.
Pollard, J., Valkovics, E. (1992). The Gompertz distribution and its applications. *Genus* 48:15–29.
Preston, S., Heuveline, P., Guillot, M. (2001). *Demography: Measuring and Modeling Population Processes*. Oxford: Blackwell.
Promislow, D. (1991). Senescence in natural populations of mammals: a comparative study. *Evolution*. (8):1869–1887.
Ricklefs, R., Scheuerlein, A. (2002). Biological implications of the Weibull and Gompertz models of aging. *J. Gerontol. Ser. A: Biol. Sci. Med. Sci.* 57:B69–B76.
Sen, P. (1959). On the moments of the sample quantiles. *Calcutta Stat. Assoc. Bull.* 9:1–19.
Shin, H., Jung, Y., Jeong, C., Heo, J. (2011). Assessment of modified Anderson–Darling test statistics for the generalized extreme value and generalized logistic distributions. *Stochastic Environ. Res. Risk Assess.* 26:105–114.
Sinclair, C., Spurr, B., Ahmad, M. (1990). Modified Anderson Darling test. *Commun. Stat.-Theory Methods* 19:3677–3686.
Song, K. (2002). Goodness-of-fit tests based on Kullback-Leibler discrimination information. *IEEE Trans. Inf. Theory* 48:1103–1117.
Stephens, M. (1974). EDF statistics for goodness of fit and some comparisons. *J. Am. Stat. Assoc.* 69:730–737.
Stephens, M. (1977). Goodness of fit for the extreme value distribution. *Biometrika* 64:583–588.
Strehler, B. L., Mildvan, A. S. (1960). General theory of mortality and aging. a stochastic model relates observations on aging, physiologic decline, mortality, and radiation. *Science* 132(3418):14–21.
Thatcher, A. (1999). The long-term pattern of adult mortality and the highest attained age. *J. R. Stat. Soc.: Ser. A (Stat. Soc.)* 162:5–43.
Vaupel, J. (1986). How change in age-specific mortality affects life expectancy. *Popul. Stud.* 40:147–157.
Vaupel, J., Manton, K., Stallard, E. (1979). The impact of heterogeneity in individual frailty on the dynamics of mortality. *Demography* 16:439–454.
Véron, J., Rohrbasser, J. (2003). Wilhelm lexis: the normal length of life as an expression of the nature of things (English edition). *Population* 58:303–322.

Vogel, R. (1986). The probability plot correlation coefficient test for the normal, lognormal, and gumbel distributional hypotheses. *Water Resources Res.* 22:587–590.

Watterson, I., Dix, M. (2003). Simulated changes due to global warming in daily precipitation means and extremes and their interpretation using the gamma distribution. *J. Geophys. Res* 108:3-1–3-20.

Wilks, S. (1938). The large-sample distribution of the likelihood ratio for testing composite hypotheses. *Ann. Math. Stat.* 9:60–62.

Willekens, F. (2001). Gompertz in context: the Gompertz and related distributions. In: Tabeau, E., van den Berg Jeths, A., Heathcote, C., eds. *Forecasting Mortality in Developed Countries – Insights from a Statistical, Demographic and Epidemiological Perspective, European Studies of Population*(Vol. 9, pp. 105–126). Dordrecht (Netherlands): Kluwer Academic Publishers.

Willemse, W., Koppelaar, H. (2000). Knowledge elicitation of Gompertz' law of mortality. *Scand. Actuarial J.* 2000:168–179.

Wilson, D. (1994). The analysis of survival (mortality) data: fitting Gompertz, Weibull, and logistic functions. *Mech. Ageing Dev.* 74:15–33.

Witten, M., Satzer, W. (1992). Gompertz survival model parameters: estimation and sensitivity. *Appl. Math. Lett.* 5:7–12.