Statistical properties and different methods of estimation of Gompertz distribution with application

Sanku Dey, Fernando A. Moala & Devendra Kumar

Taylor & Francis
Taylor & Francis Group

Statistical properties and different methods of estimation of Gompertz distribution with application

Sanku Dey
Department of Statistics
St. Anthony's College
Shillong 793001
Meghalaya
India

Fernando A. Moala
Department of Statistics
State University of Sao Paulo
Sao Paulo
Brazil

Devendra Kumar*
Department of Statistics
Central University of Haryana
Mahendergarh 123031
Haryana
India

Abstract

This article addresses the various properties and different methods of estimation of the unknown parameters of Gompertz distribution. Although, our main focus is on estimation from both frequentist and Bayesian point of view, yet, various mathematical and statistical properties of the Gompertz distribution (such as quantiles, moments, moment generating function, hazard rate, mean residual lifetime, mean past lifetime, stochasic ordering, stress-strength parameter, various entropies, Bonferroni and Lorenz curves and order statistics) are derived. We briefly describe different frequentist approaches, namely, maximum likelihood estimators, moments estimators, pseudo-moments estimators, modified moments estimators, L-moment estimators, percentile based estimators, least squares and weighted least squares estimators, maximum product of spacings estimators, minimum spacing

E-mail: `devendrastats@gmail.com`

absolute distance estimators, minimum spacing absolute-log distance estimator, Cramér-von-Mises estimators, Anderson-Darling and right-tail Anderson-Darling and compare them using extensive numerical simulations. Coverage probabilities for the frequentist methods are also obtained. Next we consider Bayes estimation under different types of loss function (symmetric and asymmetric loss functions) using gamma priors for both shape and scale parameters. Furthermore, the Bayes estimators and their respective posterior risks are computed and compared using MCMC algorithm. Finally, a real data set have been analyzed for illustrative purposes.

1. Introduction

Benjamin Gompertz (1825) introduced the Gompertz distribution in connection with human mortality and actuarial Tables. Since then, considerable attention has received from demographers and actuaries. This distribution is a generalization of the exponential distribution and has many real life applications, especially in medical and actuarial studies. It has some nice relations with some of the well-known distributions such as exponential, double exponential, Weibull, extreme value (Gumbel Distribution) or generalized logistic distribution (see Willekens (2002)). An important characteristic of the Gompertz distribution is that it has an exponentially increasing failure rate for the life of the systems and is often used to model highly negatively skewed data in survival analysis (Johnson and Johnson (1979)). In recent past, many authors have contributed to the studies of statistical methodology and characterization of this distribution; for example, Read (1983), Makany (1991), Rao and Damaraju (1992), Franses (1994), Chen (1997) and Wu and Lee (1999). Garg et al. (1970) studied some of the statistical properties of the Gompertz distribution and estimated the parameters using maximum likelihood method with reference to the data on the effects of prolonged oral contraception on mortality in mice. Jaheen (2003a, 2003b) studied this distribution based on progressive type-II censoring and record values using Bayesian approach. Wu et al. (2003) derived the point and interval estimators for the parameters of the Gompertz distribution based on progressive type II censored samples. Wu et al. (2004) used least squared method to estimate the parameters of the Gompertz distribution. Ismail (2010) discussed the point and interval estimations of a two-parameter Gompertz distribution under partially accelerated life tests with Type-II censoring. Kiani et al. (2012) studied

the performance of the Gompertz model with time-dependent covariate in the presence of right censored data. Moreover, they compared the performance of the model under different censoring proportions (CP) and sample sizes. Recently, Lenart (2014) obtained moments of the Gompertz distribution and maximum likelihood estimators of its parameters.

We have a stream of estimation methods available for the parametric distribution in the literature, some of the estimation methods are well researched on theoretical aspect. However, it is worth noting that in the case of small samples, there is often evidence that the maximum likelihood method does not perform well. Therefore, other estimating methods have recently been developed. The appeal of the methods of estimation vary from user to user and area of application. For instance, one may prefer to use the uniformly minimum variance estimation method although the estimator does not have a closed form expression.

The motivation of the paper is two fold: first is to study the properties of the Gompertz distribution (GM), and second is to estimate the parameters of the model from both frequentist and Bayesian viewpoints for different sample sizes and different parameter values and to develop a guideline for choosing the best estimation method for the Gompertz distribution, which we think would be of deep interest to applied statisticians.

The uniqueness of this study comes from the fact that thus far, no attempt has been made to compare all these estimators for the two-parameter Gompertz distribution along with statistical properties. Different estimation methods were compared for generalized Rayleigh distributions by Kundu and Raqab (2005); for generalized logistic distributions by Alkasasbeh and Raqab (2009); for the Weibull distribution by Teimouri et al. (2013) and Dey et al. (2014, 2015, 2016, 2017a, 2017b, 2017c, 2017d, 2017e, 2017f) for the two-parameter Rayleigh, weighted exponential, two-parameter Maxwell, exponentiated-Chen, Dagum, transmuted-Rayleigh, two parameter exponentiated-Gumbel, new extension of generalized exponential and NH distributions. Maximum likelihood estimates using different experimental schemes studied by Azzam (1994) and asymptotic normality of maximum likelihood estimators for non-parametric Markov chains was studied by AL-Eideh (1996). The paper is organized as follows. Various mathematical and statistical properties of the Gompertz distribution are presented in Section 2. Section 3 describes fourteen frequentist methods of estimation. The Bayes estimators are presented in section 4. In Section 5, performance of several estimation procedures along with coverage percentages of bootstrap confidence interval using frequentist approaches are provided. The methodology developed in this

manuscript and the usefulness of the Gompertz distribution is illustrated by using a real data example. Finally, concluding remarks are provided in Section 6.

A random variable X has the Gompertz distribution with parameters c and λ, say $GM(c, \lambda)$, if its density function is

$$f(x;\lambda,c) \;=\; \lambda e^{cx} e^{-\frac{\lambda}{c}(e^{cx}-1)}; \qquad x>0, c, \lambda > 0, \tag{1}$$

and the corresponding c.d.f is given by

$$F(x;\lambda,c) \;=\; 1 - e^{-\frac{\lambda}{c}(e^{cx}-1)}; \qquad x > 0, c, \lambda > 0. \tag{2}$$

It can be seen that

$$f(x;\lambda,c) \;=\; [\lambda + c(-\log[\bar{F}(x)])]\bar{F}(x); \quad x>0, c, \lambda > 0,$$

where $\bar{F}(x) = 1 - F(x)$. Below, we state and prove a theorem which characterizes the distribution:

Theorem 1: *The random variable X follows a Gompertz distribution with parameters λ and c if and only if the density function f satisfies the homogeneous differential equation of the form:*

$$f' + (\lambda e^{cx} - c)f \;=\; 0 \tag{3}$$

where prime denotes first order differentiation.

Proof. Suppose X is a Gompertz distribution random variable, then $f(x; \lambda, c)$ and $f'(x; \lambda, c)$ are the pdf and the first derivative of the pdf of the Gompertz distribution. By substituting $f(x)$ and $f'(x)$ in the differential equation (3), the equation is satisfied.

Conversely, we assume that f satisfies equation (3), then we have

$$\int \frac{f'}{f} dx \;=\; -\lambda \int e^{cx} dx + c \int dx. \tag{4}$$

After simplification we get,

$$f \;=\; D e^{cx} e^{-\frac{\lambda}{c}(e^{cx}-1)}, \qquad x > 0 \tag{5}$$

where D is the normalizing constant and the value of $D = \lambda$. $\square$

Application of the Theorem: From the homogeneous differential equation (3), we get

$$x = \frac{1}{c}\log\frac{(cf + f')}{\lambda f} \tag{6}$$

or equivalently,

$$x = \frac{1}{c}\log\left(\frac{cF' + F''}{\lambda F'}\right) \tag{7}$$

where F is the corresponding cdf of the Gompertz distribution. Thus, the importance of this theorem lies in the linearizing transformation (6) and (7) which could be regarded as Gompertz model alternative to the Berkson's logit transform (Berkson (1944)) for the ordinary logistic model and Ojo (1997) logit transform for generalized logistic model. Hence equation (6) or (7) could be referred as Gompertz logit transform. Thus the theorem shows the flexibility of Gompertz distribution.

2. Statistical and Mathematical Properties of Gompertz Distribution

The quantile function $x_p = Q(p) = F^{-1}(p)$, for $0 < p < 1$, of the Gompertz distribution is obtained from (2), it follows that the quantile function x_p is

$$x_p = \frac{1}{c}\ln\left[1 - \frac{c}{\lambda}\ln(1 - p)\right]. \tag{8}$$

In particular, the median of the Gompertz distribution can be written as

$$Md(X) = M_d = \frac{1}{c}\ln\left[1 - \frac{c}{\lambda}\ln(1 - 0.5)\right]. \tag{9}$$

Figure 1(a) shows the plot of the pdf of the Gompertz distribution for different values of the parameters c and λ and from the plot, it is quite evident that the Gompertz distribution is positively skewed distribution.

If the random variable X is distributed $GM(c, \lambda)$, then its nth moment around zero can be expressed as

$$E(X^n) = \frac{\lambda e^{\frac{\lambda}{c}}}{c}\int_1^\infty \frac{1}{c^n}e^{-\frac{\lambda}{c}x}[\ln(x)]^n \, dx. \tag{10}$$

On simplification, we get

$$E(X^n) = \frac{n!}{c^n}e^{\frac{\lambda}{c}}E_1^{n-1}\left(\frac{\lambda}{c}\right), \tag{11}$$

where

$$E_s^n(z) = \frac{1}{n!}\int_1^\infty (\ln(x))^n x^{-s} e^{-zx} dx, \tag{12}$$

$$E_n(x) = \int_1^\infty \frac{e^{-xt}}{t^n} dt,$$

and

$$E_s^0(z) = E_s(z),$$

is the generalized integro-exponential function (Milgram(1985)). The variance, skewness and kurtosis measures can now be calculated using the relations

$$Var(X) = E(X^2) - E^2(X),$$

$$Skewness(X) = \frac{E(X^3) - 3E(X)E(X^2) + 2E^3(X)}{Var^{3/2}(X)},$$

$$Kurtosis(X) = \frac{E(X^4) - 4E(X)E(X^3) + 6E(X^2)E^2(X) - 3E^4(X)}{Var^2(X)}.$$

It can be noticed from Table 1 that both the mean and the variance of the Gompertz distribution are increasing function of λ where as both the skewness and kurtosis are decreasing function of λ (see Fig.2).

2.1 *Moment Generating Function*

Many of the interesting characteristics and features of a distribution can be obtained via its moment generating function and moments. Let X denote a random variable with the probability density function (1). By definition of moment generating function of X and using (1), we have

$$M_x(t) = E(e^{tx}) = \int_0^\infty e^{tx} f(x) dx = \frac{\lambda}{c} \sum_{p=0}^\infty (-1)^p \binom{\frac{t}{\lambda}}{p} \frac{\Gamma(p+1)}{\left(\frac{\lambda}{c}\right)^{p+1}}. \tag{12}$$

2.2 *Hazard function*

The basic tools for studying the ageing and reliability characteristics of the system are the hazard rate (HR) and the mean residual lifetime

(MRL). Thus the hazard rate function of the Gompertz distribution is given by

$$H(x;\lambda,c) \;=\; \frac{f(x;\lambda,c)}{1-\lambda(x;\lambda,c)} = \lambda e^{cx}. \tag{13}$$

Note that the hazard rate function is increasing function if $c > 0$ or constant function if $c = 0$. Figure 1(b) shows the shapes of the hazard function for different selected values of the parameters c and λ. From the plot, it is quite evident that the Gompertz distribution has increasing hazard rate function.

2.3 *Mean residual Life Function*

The mean residual life (MRL) is the expected remaining life, $X - x$, given that the item has survived to time x. Thus, in life testing situations, the expected additional lifetime given that a component has survived until time x is called the (MRL). Since the (MRL) function is the expected remaining life, x must be subtracted, yielding for $x > 0$.

$$\mu(x) \;=\; E(X-x\,|\,X>x)=\frac{\int_x^{\infty}S(t)dt}{S(t)}$$

$$\Rightarrow \qquad \int_x^{\infty}S(t)dt \;=\; \int_x^{\infty}exp\left[\frac{\lambda}{c}(1-e^{ct})\right]dt$$

let $z=\frac{\lambda}{c}e^{ct}$, we have

$$\int_x^{\infty}S(t)dt \;=\; \frac{e^{\frac{\lambda}{c}}}{c}\Gamma_{inc}(0,z),$$

where

$$\Gamma_{inc}(a,x) \;=\; \int_x^{\infty}t^{a-1}e^{-t}dt, x,a \geq 0.$$

Therefore,

$$\mu(x) \;=\; \frac{e^{z}}{c}\Gamma_{inc}(0,z). \tag{14}$$

2.4 *Mean Past Lifetime (MPL)*

Mean past lifetime function is an important reliability measure for evaluating systems' performance. Assume that a component with lifetime

X has failed before time x, i.e., $X \geq x$. Hence, the mean past lifetime (*MPL*) of the component can be defined by $k(x)$, i.e., the expectation of the conditional random variable $x - X \mid X \geq x$ as

$$k(x) \;=\; E(X - x \mid X \leq x) = \frac{\int_0^x F(t)dt}{F(t)} = \frac{e^{\frac{\lambda}{c}} \, \gamma_{inc}(0, \frac{\lambda}{c} e^{cx})}{c\left[1 - \exp\left\{\frac{\lambda}{c}(e^{cx} - 1)\right\}\right]},$$

where

$$\gamma_{inc}(a, x) \;=\; \int_0^x t^{a-1} e^{-t} dt, x, a \geq 0.$$

2.5 *Stochastic ordering*

Stochastic ordering has been recognized as an important tool in reliability theory and other fields to assess comparative behavior. If X and Y are independent random variables with cdfs, sfs and pdfs $F_X(x)$ and $F_Y(x)$, $\bar{F}_X(x) = 1 - F_X(x)$ and $\bar{F}_Y(x) = 1 - F_Y(x)$, and $f_x(x)$ and $f_y(x)$, respectively, then X is said to be smaller than Y in the following ordering as:

- stochastic order ($X \leq_{st} (Y)$) if $F_X(x) \geq F_Y(x)$ for all x
- hazard rate order ($X \leq_{hr} (Y)$) if $h_X(x) \geq h_Y(x)$ for all x
- mean residual life order ($X \leq_{mrl} (Y)$) if $m_X(x) \geq m_Y(x)$ for all x
- likelihood ratio order ($X \leq_{lr} (Y)$) if $\frac{f_X(x)}{f_Y(x)}$ decrease in x.

The following theorem reveals that the Gompertz distributions are ordered with respect to strongest likelihood ratio ordering when appropriate assumptions are satisfied.

Theorem 2: *Let $X \sim G(c_1, \lambda_1)$ and $Y \sim G(c_2, \lambda_2)$. If $c_1 = c_1 = c$ and $\lambda_1 \geq \lambda_2$, then $X \leq_{lr} Y, X \leq_{hr} Y, X \leq_{mrl} Y$ and $X \leq_{st} Y$.*

Proof. The likelihood ratio is

$$\frac{f_X(x)}{f_Y(x)} \;=\; \frac{\lambda_1 \, e^{c_1 x} \exp\left[-\frac{\lambda_1}{c_1}(e^{c_1 x} - 1)\right]}{\lambda_2 \, e^{c_2 x} \exp\left[-\frac{\lambda_2}{c_2}(e^{c_2 x} - 1)\right]}$$

thus,

$$\frac{d}{dx}\log\frac{f_X(x)}{f_Y(x)} = (c_1-c_2)-\lambda_1\,e^{c_1 x}+\lambda_2\,e^{c_2 x}.$$

Now if $c_1 = c_2 = c$ and $\lambda_1 \geq \lambda_2$ then $\dfrac{d}{dx}\log\dfrac{f_X(x)}{f_Y(x)} \leq 0,$ which implies that $X \leq_{lr} Y$ and hence $X \leq_{lr} Y$, $X \leq_{hr} Y$, $X \leq_{mrl} Y$ and $X \leq_{st} Y$.

2.6 *Entropies*

The entropy of a random variable measures the variation of the uncertainty. A large value of entropy indicates the greater uncertainty in the data. Some popular entropy measures are Rényi entropy (Rényi, 1961) or Shannon entropy (Shannon, 1951). If X is a non-negative continuous random variable with *pdf* $f(x)$ then Shannon's entropy is defined as:

$$H(f) = E[-\ln f(x)]=-\int_0^{\infty}f(x)\ln f(x)dx.$$

and this is usually referred to as the continuous entropy (or differential entropy). An explicit expression of Shannon entropy for Gompertz distribution is obtained as

$$H(f) = -\ln\lambda-\frac{\lambda}{c}-\lambda c\sum_{p=0}^{\infty}\frac{(-1)^p\,e^{\frac{\lambda}{c}}}{\Gamma(p+1)}\left(\frac{\lambda}{c}\right)^p\left(-\frac{1}{2c}\right)^{p+1}$$
$$-\frac{\lambda}{c}\sum_{p=0}^{\infty}(-1)^p\binom{\frac{t}{c}}{p}\frac{\Gamma(p+1)}{\left(\frac{\lambda}{c}\right)^{p+1}}. \tag{15}$$

If X has the probability distribution function $f(\cdot)$, Rényi entropy (Rényi, 1961) can be expressed as

$$H_{\delta}(x) = \frac{1}{1-\delta}\ln\left(\int_0^{\infty}f^{\delta}(x)dx\right), \qquad \delta>0, \qquad \delta\neq 1,$$
$$= -\ln\delta+\frac{1}{1-\delta}\ln\sum_{p=1}^{\infty}\sum_{q=0}^{\infty}(-1)^q\binom{\delta-1}{p}\binom{p}{q}\left(\frac{c}{\lambda}\right)^p\frac{\Gamma(p+1)}{(p+1)^{q-1}}. \tag{16}$$

2.7 *Bonferroni and Lorenz curve*

Boneferroni and Lorenz curves are proposed by Bonferroni (1930). These curves have applications not only in economics to study income and

poverty, but also in other fields like reliability, demography, insurance and medicine. They are define as

$$B(p) = \frac{1}{p\mu}\int_0^q xf(x)dx \qquad (17)$$

and

$$L(p) = \frac{1}{\mu}\int_0^q xf(x)dx, \qquad (18)$$

respectively, where $\mu = E(X)$ and $q = F^{-1}(p)$. By using (1), one can reduce (18) and (19) to

$$B(p) = \frac{\lambda\, e^{\lambda/c}}{p\,\mu}\sum_{k=0}^{\infty}\sum_{l=0}^{\infty}\frac{(-1)^k\left(\dfrac{\lambda}{c}\right)^k [c(k+1)]^l q^{l+2}}{k!\,l!\,(l+2)},$$

$$L(p) = \frac{\lambda\, e^{\lambda/c}}{p}\sum_{k=0}^{\infty}\sum_{l=0}^{\infty}\frac{(-1)^k\left(\dfrac{\lambda}{c}\right)^k [c(k+1)]^l q^{l+2}}{k!\,\,l!\,(l+2)},$$

respectively.

2.8 *Order Statistics*

Moments of order statistics play an important role in quality control testing and reliability to predict the failure of future items based on the times of few early failures. We know that if $X_{(1)} \le \cdots \le X_{(n)}$ denotes the order statistic of a random sample $X_1, \cdots, X_n$ from a continuous population with cdf $F(x)$ and pdf $f(x)$ then the pdf of $X_{(j)}$ is given by

$$f_{X_{(j)}}(x) = \frac{n!}{(j-1)!(n-j)!}f(x)\left(F(x)\right)^{j-1}\left(1-F(x)\right)^{n-j}, \qquad (19)$$

for $j = 1, \ldots, n$. The *pdf* and *cdf* of the jth order statistic for a Gompertz distribution is given by

$$\begin{aligned} f_{X_{(j)}}(x) = {} & \frac{\lambda\, n!}{(j-1)!(n-j)!}\sum_{u=0}^{j-1}(-1)^u\binom{j-1}{u}e^{cx} \\ & \exp\left[-\frac{(u+n-j+1)\lambda}{c}(e^{cx}-1)\right] \end{aligned} \qquad (20)$$

and

$$F_{j:n}(x) = \sum_{l=k}^{n}\sum_{u=0}^{j-1}(-1)^u \binom{n}{l}\binom{j-1}{u}\exp\left[-\frac{(n+l)\lambda}{c}(e^{cx}-1)\right].$$ (21)

The $k - th$ moments of $X_{j:n}$ can be expressed

$$E[X_{j:n}^k] = \frac{\lambda\,\Gamma(k+1)n!}{(j-1)!(n-j)!}\sum_{p=0}^{\infty}\sum_{u=0}^{j-1}\frac{(-1)^{u+p}}{\Gamma(p+1)}\binom{j-1}{u}$$ (22)

$$\times\exp\left[\frac{(u+n-j+1)\lambda}{c}\right]\left[\frac{(u+n-j+1)\lambda}{c}\right]^p\left(-\frac{1}{c(k+1)}\right)^{p+1}.$$

An application of the first moments of order statistics can be considered in calculating the L-moments which are in fact the linear combinations of the expected order statistics. See Hosking (1990) for details.

2.9 *Stress Strength Parameter*

Here, we derive the reliability $R = Pr(X_1 > X_2)$ when X_1 and X_2 are independent Gompertz random variables distributed with parameters (c_1, λ_1) and (c_2, λ_2), then

$$R = P(X_2 < X_1) = \int_0^{\infty} f_1(x)F_2(x)dx$$

$$= 1-\exp\left\{\frac{\lambda_1}{c_1}+\frac{\lambda_1}{c_1}\right\}\sum_{p=0}^{\infty}\frac{(-1)^p\left(\frac{\lambda_2}{c_2}\right)^p\left(\frac{c_1}{\lambda_1}\right)^{pc_2/\lambda_1}}{p!}\left[\Gamma\left(p\frac{c_2}{c_1}+1\right)\right]$$

$$\sum_{q=0}^{\infty}\frac{(-1)^q\left(\frac{\lambda_1}{c_1}\right)^{\left(\frac{c_2}{c_1}\right)p+q+1}}{q!\left[\left(\frac{c_2}{c_1}\right)p+q+1\right]}.$$

If $c_1 = c_2 = c$, then

$$R = \frac{\lambda_2}{\lambda_1+\lambda_2}.$$

3. Methods of Estimation

3.1 *Method of Maximum Likelihood*

The method of maximum likelihood is the most frequently used method of parameter estimation (Casella and Berger(1990)). The success of the method stems no doubt from its many desirable properties including consistency, asymptotic efficiency, invariance property as well as its intuitive appeal. Let $x_1, \cdots, x_n$ be a random sample of size n from (1), then the log-likelihood function of (1) without constant terms is given by

$$\ell(c, \lambda; x) \;=\; \log L(c, \lambda; x) = n\log\lambda + c\sum_{i=1}^{n} x_i - \frac{\lambda}{c}\sum_{i=1}^{n}(e^{cx_i} - 1).$$

For ease of notation, we will denote the first partial derivatives of any function $f(x, y)$ by f_x and f_y. Now setting

$$\ell_c \;=\; 0 \qquad \text{and} \qquad \ell_\lambda = 0,$$

we have

$$\ell_\lambda \;=\; \frac{n}{\lambda} - \frac{1}{c}\sum_{i=1}^{n}(e^{cx_i} - 1) = 0, \tag{24}$$

and

$$\ell_c \;=\; \sum_{i=1}^{n} x_i - \frac{\lambda}{c}\sum_{i=1}^{n} x_i e^{cx_i} + \frac{\lambda}{c^2}\sum_{i=1}^{n} e^{cx_i} - \frac{n\lambda}{c^2} = 0. \tag{25}$$

The MLEs $\hat{c}$ and $\hat{\lambda}$ of c and λ are obtained by solving this nonlinear system of equations. Since the MLE of the vector of unknown parameters $\theta = (c, \lambda)$ cannot be derived in closed forms, therefore it is not easy to derive the exact distributions of the MLEs and hence we can not get the exact bounds of the parameters. The idea is to use the large sample approximation. It is known that the asymptotic distribution of the MLE $\hat{\theta}$ is

$$(\hat{\theta} - \theta) \to N_2(0, I^{-1}(\theta))$$

(see Lawless (1982)), where $I^{-1}(\theta)$ is the inverse of the observed information matrix of the unknown parameters $\theta = (c, \lambda)$.

$$I^{-1}(\theta) \;=\; \left(\begin{array}{cc} -\dfrac{\partial^2 \log L}{\partial c^2} & -\dfrac{\partial^2 \log L}{\partial c \partial \lambda} \\[2ex] -\dfrac{\partial^2 \log L}{\partial \lambda \partial c} & -\dfrac{\partial^2 \log L}{\partial \lambda^2} \end{array} \right)^{-1}_{\Big|(c,\lambda)=(\hat{c},\hat{\lambda})}$$

$$= \begin{pmatrix} \text{var}(\hat{c}_{MLE}) & \text{cov}(\hat{c}_{MLE}, \hat{\lambda}_{MLE}) \\ \text{cov}(\hat{\lambda}_{MLE}, \hat{c}_{MLE}) & \text{var}(\hat{\lambda}_{MLE}) \end{pmatrix} = \begin{pmatrix} \sigma_{cc} & \sigma_{c\lambda} \\ \sigma_{\lambda c} & \sigma_{\lambda\lambda} \end{pmatrix}.$$

The derivatives in $I(\theta)$ are given as follows

$$\left. \frac{\partial^2 \log L}{\partial c^2} \right|_{c=\hat{c}_{MLE}} = \frac{\lambda}{c^2} \left[-c\sum_{i=1}^{n} x_i^2 e^{cx_i} + 2\sum_{i=1}^{n} x_i e^{cx_i} - \frac{2}{c}\sum_{i=1}^{n} e^{cx_i} + \frac{2n}{c} \right],$$

$$\left. \frac{\partial^2 \log L}{\partial \lambda^2} \right|_{\lambda=\hat{\lambda}_{MLE}} = -\frac{n}{\lambda^2},$$

$$\left. \frac{\partial^2 \log L}{\partial c \partial \lambda} \right|_{c=\hat{c}_{MLE}, \lambda=\hat{\lambda}_{MLE}} = \left[\frac{1}{c^2}\sum_{i=1}^{n}(e^{cx_i} - 1) - \frac{1}{c}\sum_{i=1}^{n} x_i e^{cx_i} \right].$$

Therefore, the above approach is used to derive the approximate $100(1 - \tau)\%$ confidence intervals of the parameters $\theta = (c, \lambda)$ as in the following forms

$$\hat{c}_{MLE} \pm z_{\frac{\tau}{2}} \sqrt{Var(\hat{c}_{MLE})}, \hat{\lambda}_{MLE} \pm z_{\frac{\tau}{2}} \sqrt{Var(\hat{\lambda}_{MLE})}.$$

Here, $Z_{\frac{\tau}{2}}$ is the upper $\left(\frac{\tau}{2}\right)$th percentile of the standard normal distribution.

3.2 *Method of Moments and pseudo-moments*

The MMEs of the Gompertz distribution can be obtained by equating the first two theoretical moments of (1) with the sample moments $m_1 = \frac{1}{n}\sum_{i=1}^{n} x_i$ and $m_2 = \frac{1}{n}\sum_{i=1}^{n} x_i^2$ respectively. Thus the MMEs $\hat{c}$ and $\hat{\lambda}$ of c and λ are

$$\frac{1}{n}\sum_{i=1}^{n} x_i = \frac{1}{c} e^{\frac{\lambda}{c}} E_1^0\left(\frac{\lambda}{c}\right)$$

and

$$\frac{1}{n}\sum_{i=1}^{n} x_i^2 = \frac{2!}{c^2} e^{\frac{\lambda}{c}} E_1^1\left(\frac{\lambda}{c}\right).$$

Based on the sample moments m_1 and m_2, the pseudo-moment estimators (PME) $\tilde{c}$ and $\tilde{\lambda}$ of c and λ are obatined by

$$S(X,\tilde{c},\tilde{\lambda}) = \min_{c,\lambda} S(X,c,\lambda)$$

where

$$S(X,c,\lambda) = \left[m_1 - E(x) \right]^2 + \left[m_2 - E(x^2) \right]^2.$$

3.3 *Method of Modified Moments*

Another way of finding the MME is using the sample mean and the sample variance. We know that

$$E(X \mid c,\lambda) = \frac{1}{c} e^{\frac{\lambda}{c}} E_1^0 \left(\frac{\lambda}{c} \right) \tag{26}$$

and

$$Var(X \mid c,\lambda) = \frac{2!}{c^2} e^{\frac{\lambda}{c}} E_1^1 \left(\frac{\lambda}{c} \right) - \left[\frac{1}{c} e^{\frac{\lambda}{c}} E_1^0 \left(\frac{\lambda}{c} \right) \right]^2.$$

The population coefficient of variation is

$$CV(X \mid c,\lambda) = \frac{\sqrt{Var(X \mid c,\lambda)}}{E(X \mid c,\lambda)}$$

which is independent of the parameter c. So, the estimate $\hat{\lambda}_{MMME}$ for λ can be obtained by solving the non-linear equation

$$\frac{\sqrt{Var(X \mid c,\lambda)}}{E(X \mid c,\lambda)} - \frac{s}{\bar{x}} = 0 \tag{27}$$

where $\bar{x}$ and s are the sample mean and sample standard deviation respectively. Once we have $\hat{\lambda}_{MMME}$ for λ, we can obtain $\hat{c}_{MMME}$.

3.4 *Method of Least-Squares and Weighted least squares*

The least square estimators and weighted least square estimators were proposed by Swain et al. (1988) to estimate the parameters of Beta distributions. Suppose $F(X_{i:n})$ denotes the distribution function of the ordered random variables $X_{1:n} < X_{2:n} < \cdots < X_{n:n}$ where $\{X_1, X_2, \cdots, X_n\}$ is a random sample of size n from a distribution function $F(\cdot)$. Therefore, in

this case, the least square estimators of c and λ, say $\hat{c}_{LSE}$ and $\hat{\lambda}_{LSE}$ can be obtained by minimizing

$$S(c,\lambda) \;=\; \sum_{i=1}^{n}\left[F\left(x_{i:n}\,|\,c,\lambda\right)-\frac{i}{n+1}\right]^{2}$$

with respect to c and λ, where $F(\cdot)$ is the cdf of the Gompertz distribution. Equivalently, they can be obtained by solving:

$$\sum_{i=1}^{n}\left[F\left(x_{i:n}\,|\,c,\lambda\right)-\frac{i}{n+1}\right]\eta_{1}\left(x_{i:n}\,|\,c,\lambda\right)=0$$

$$\sum_{i=1}^{n}\left[F\left(x_{i:n}\,|\,c,\lambda\right)-\frac{i}{n+1}\right]\eta_{2}\left(x_{i:n}\,|\,c,\lambda\right)=0.$$

The WLSEs, $\hat{c}_{WLSE}$ and $\hat{\lambda}_{WLSE}$, can be obtained by minimizing

$$W(c,\lambda) \;=\; \sum_{i=1}^{n}\frac{(n+1)^{2}(n+2)}{i(n-i+1)}\left[F\left(x_{i:n}\,|\,c,\lambda\right)-\frac{i}{n+1}\right]^{2}.$$

These estimators can also be obtained by solving:

$$\sum_{i=1}^{n}\frac{(n+1)^{2}(n+2)}{i(n-i+1)}\left[F\left(x_{i:n}\,|\,c,\lambda\right)-\frac{i}{n+1}\right]\eta_{1}\left(x_{i:n}\,|\,c,\lambda\right)=0$$

$$\sum_{i=1}^{n}\frac{(n+1)^{2}(n+2)}{i(n-i+1)}\left[F\left(x_{i:n}\,|\,c,\lambda\right)-\frac{i}{n+1}\right]\eta_{2}\left(x_{i:n}\,|\,c,\lambda\right)=0$$

where

$$\eta_{1}\left(x_{i:n}\,|\,c,\lambda\right) \;=\; \frac{\lambda}{c^{2}}e^{-\frac{\lambda}{c}(e^{cx_{i:n}}-1)}(1+e^{cx_{i:n}}(cx_{i:n}-1)) \tag{28}$$

and

$$\eta_{2}\left(x_{i:n}\,|\,c,\lambda\right) \;=\; \frac{1}{c}e^{-\frac{\lambda}{c}(e^{cx_{i:n}}-1)}(e^{cx_{i:n}}-1). \tag{29}$$

3.5 *Method of Percentile*

The percentile estimator is a statistical method used to estimate the parameters by comparing the sample points with the theoretical points. This method was originally suggested by Kao (1958, 1959). Since the cumulative distribution of the Gompertz distribution is

$$F(x;\lambda,c) \;=\; 1-e^{-\frac{\lambda}{c}(e^{cx}-1)},$$

therefore, the quantile function is given by

$$x_p = \frac{1}{c}\ln[1-\frac{c}{\lambda}\ln\left(1-F(x)\right)].$$

Let $X_{(j)}$ be the jth order statistic, i.e, $X_{(1)} \le X_{(2)} \le \cdots \le X_{(n)}$. If p_j denotes some estimate of $F(x_{(j)};c,\lambda)$, then the estimators of c and λ can be obtained by minimizing

$$\sum_{j=1}^{n}\left(x_{(j)} - \frac{1}{c}\ln[1-\frac{c}{\lambda}\ln(1-p_j)]\right)^2$$

with respect to c and λ. The estimators of c and λ can be obtained by solving the following nonlinear equations

$$\sum_{j=1}^{n}\left(x_{(j)} - \frac{1}{c}\ln[1-\frac{c}{\lambda}\ln(1-p_j)]\right)\left[\frac{\frac{c}{\lambda}\ln(1-p_j)-\left(1-\frac{c}{\lambda}\ln(1-p_j)\right)\ln\left(1-\frac{c}{\lambda}\ln(1-p_j)\right)}{c^2\left(1-\frac{c}{\lambda}\ln(1-p_j)\right)}\right] = 0$$

$$\sum_{j=1}^{n}\left(x_{(j)} - \frac{1}{c}\ln[1-\frac{c}{\lambda}\ln(1-p_j)]\right)\left[\frac{\frac{1}{\lambda^2}\ln(1-p_j)}{\left(1-\frac{c}{\lambda}\ln(1-p_j)\right)}\right] = 0$$

respectively. We call the corresponding estimators as the percentile estimators or PCE's. Several estimators of p_j can be used here, see for example Mann et al. (1974). In this paper, we consider $p_j = \frac{j}{n+1}$.

3.6 *Method of L-Moments*

In this section, we provide the L-moments estimators which can be obtained by equating the first two sample L-moments with the corresponding population L-moments. The first two sample L-moments are

$$l_1 = \frac{1}{n}\sum_{i=1}^{n}x_{(i)}, \qquad l_2 = \frac{2}{n(n-1)}\sum_{i=1}^{n}(i-1)x_{(i)} - l_1$$

and the first two population L-moments are

$$\lambda_1 = E(X_{1:1}) = E(X) = \frac{1}{c}e^{\frac{\lambda}{c}}E_1^0\left(\frac{\lambda}{c}\right)$$

$$\lambda_2 = \frac{1}{2}\left[E(X_{2:2}) - E(X_{2:1})\right],$$

where

$$E(X_{2:2}) = 2\int_0^\infty xf(x)F(x)dx$$

On simplification, we get

$$= 2\frac{1}{c}e^{\frac{\lambda}{c}}E_1^0\left(\frac{\lambda}{c}\right) - \frac{1}{c}e^{\frac{2\lambda}{c}}E_1^0\left(\frac{2\lambda}{c}\right)$$

and

$$E(X_{1:2}) = 2\int_0^\infty xf(x)dx - 2\int_0^\infty xf(x)F(x)dx = \frac{1}{c}e^{\frac{2\lambda}{c}}E_1^0\left(\frac{2\lambda}{c}\right)$$

The L-moments estimators $\hat{c}_{LME}$ and $\hat{\lambda}_{LME}$ of the parameters c and λ can be obtained by solving numerically the following equations:

$$\lambda_1\left(\hat{c}_{LME}, \hat{l}_{LME}\right) = l_1, l_2\left(\hat{c}_{LME}, \hat{l}_{LME}\right) = l_2. \tag{30}$$

3.7 Method of Maximum Product of Spacings and Minimum Spacing Distance

Cheng and Amin (1979,1983) introduced the maximum product of spacings (MPS) method as an alternative to MLE for the estimation of parameters of continuous univariate distributions. Ranneby (1984) independently developed the same method as an approximation to the Kullback-Leibler measure of information. Using the same notations in subsection 3.4, define the uniform spacings of a random sample from the Gompertz distribution as:

$$D_i(c,\lambda) = F\left(x_{i:n} \mid c,\lambda\right) - F\left(x_{i-1:n} \mid c,\lambda\right), \qquad i = 1,2,\ldots,n,$$

where $F(x_{0:n} \mid c,\lambda) = 0$ and $F(x_{n+1:n} \mid c,\lambda) = 1$. Clearly $\sum_{i=1}^{n+1}D_i(c,\lambda) = 1$. Following Cheng and Amin (1983)), the maximum product of spacings estimates $\hat{c}_{MPS}$ and $\hat{\lambda}_{MPS}$, of the parameters c and λ are obtained by maximizing, with respect to c and λ, the geometric mean of the spacings:

$$G(c,\lambda) = \left[\prod_{i=1}^{n+1}D_i(c,\lambda)\right]^{\frac{1}{n+1}}, \tag{31}$$

or, equivalently, by maximizing the function

$$H(c,\lambda) = \frac{1}{n+1}\sum_{i=1}^{n+1}\log D_i(c,\lambda). \tag{32}$$

The estimates $\hat{c}_{MPS}$ and $\hat{\lambda}_{MPS}$ of the parameters c and λ can be obtained by solving the nonlinear equations

$$\frac{\partial H(c,\lambda)}{\partial c} = \frac{1}{n+1}\sum_{i=1}^{n+1}\frac{1}{D_i(c,\lambda)}\Big[\eta_1(x_{i:n}\,|\,c,\lambda)-\eta_1(x_{i-1:n}\,|\,c,\lambda)\Big]=0 \tag{33}$$

$$\frac{\partial}{\partial \lambda} H(c,\lambda) = \frac{1}{n+1}\sum_{i=1}^{n+1}\frac{1}{D_i(c,\lambda)}\Big[\eta_2(x_{i:n}\,|\,c,\lambda)-\eta_2(x_{i-1:n}\,|\,c,\lambda)\Big]=0 \tag{34}$$

where $\eta_1(\cdot\,|\,c,\lambda)$ and $\eta_2(\cdot\,|\,c,\lambda)$ are given by (28) and (29), respectively. Cheng and Amin (1983) showed that maximizing H as a method of parameter estimation is as efficient as MLE estimation and the MPS estimators are consistent under more general conditions than the MLE estimators.

Similarly, the minimum spacing distance estimators of $\hat{c}_{MSADE}$ and $\hat{\lambda}_{MSALDE}$, of c and λ are obtained by minimizing

$$T(c,\lambda) = \sum_{i=1}^{n+1}h\left(D_i(c,\lambda),\frac{1}{n+1}\right) \tag{35}$$

where $h(x,y)$ is an appropriate distance. Some choices of $h(x,y)$ are $|x-y|$ and $|\log x - \log y|$, which are called absolute and absolute-log distance, respectively. These estimators are called minimum spacing absolute distance estimator (MSADE) and minimum spacing absolute-log distance estimator (MSALDE). This method was originally proposed by Torabi (2008). The MSADE and MSALDE of parameters c and λ can be obtained by minimizing

$$T(c,\lambda) = \sum_{i=1}^{n+1}\left|D_i(c,\lambda)-\frac{1}{n+1}\right| \tag{36}$$

and

$$T(c,\lambda) = \sum_{i=1}^{n+1}\left|\log D_i(c,\lambda)-\log\frac{1}{n+1}\right| \tag{37}$$

with respect to c and λ respectively.

The estimators $\hat{c}_{MSADE}$ and $\hat{\lambda}_{MSADE}$ of c and λ can be obtained by solving the nonlinear equations

$$\frac{\partial}{\partial c}T(c,\lambda) \;=\; \sum_{i=1}^{n+1}\frac{D_i(c,\lambda)-\dfrac{1}{n+1}}{\left|D_i(c,\lambda)-\dfrac{1}{n+1}\right|}\cdot\left[\eta_1\left(x_{i:n}\,|\,c,\lambda\right)-\eta_1\left(x_{i-1:n}\,|\,c,\lambda\right)\right]=0$$

$$\frac{\partial}{\partial \lambda}T(c,\lambda) \;=\; \sum_{i=1}^{n+1}\frac{D_i(c,\lambda)-\dfrac{1}{n+1}}{\left|D_i(c,\lambda)-\dfrac{1}{n+1}\right|}\cdot\left[\eta_2\left(x_{i:n}\,|\,c,\lambda\right)-\eta_2\left(x_{i-1:n}\,|\,c,\lambda\right)\right]=0,$$

where $D_i(c,\lambda)\neq\frac{1}{n+1}$. The estimators $\hat{c}_{MSALDE}$ and $\hat{\lambda}_{MSALDE}$ of c and λ can be obtained by solving the nonlinear equations

$$\frac{\partial}{\partial c}T(c,\lambda) \;=\; \sum_{i=1}^{n+1}\frac{\log D_i(c,\lambda)-\log\dfrac{1}{n+1}}{\left|\log D_i(c,\lambda)-\log\dfrac{1}{n+1}\right|}\cdot\frac{1}{D_i(c,\lambda)}\cdot$$
$$\left[\eta_1\left(x_{i:n}\,|\,c,\lambda\right)-\eta_1\left(x_{i-1:n}\,|\,c,\lambda\right)\right]=0$$

$$\frac{\partial}{\partial \lambda}T(c,\lambda) \;=\; \sum_{i=1}^{n+1}\frac{\log D_i(c,\lambda)-\log\dfrac{1}{n+1}}{\left|\log D_i(c,\lambda)-\log\dfrac{1}{n+1}\right|}\cdot\frac{1}{D_i(c,\lambda)}\cdot$$
$$\left[\eta_2\left(x_{i:n}\,|\,c,\lambda\right)-\eta_2\left(x_{i-1:n}\,|\,c,\lambda\right)\right]=0,$$

where $\log D_i(c,\lambda)\neq\log\frac{1}{n+1}$.

3.8 *Methods of Minimum Distances*

In this section, we will discuss three methods based on the test statistics of Cramér-von-Mises, Anderson- Darling and Right-tail Anderson-Darling. Mainly, these methods determine the values of parameters that minimize the distance between the theoretical and empirical cumulative distribution functions (see, DÁgostino (1986); Luceño (2006)).

3.8.1 Method of Cramér-von-Mises

The Cramér-von Mises estimator is a type of maximum goodness-of-fit estimators (also called minimum distance estimators) and is based on the difference between the estimate of the cumulative distribution function and the empirical distribution function. MacDonald (1971) motivated the

choice of the Cramér-von Mises statistic by providing empirical evidence that the bias of the estimator is smaller than the other goodness-of-fit estimators.

Thus, the Cramér-von Mises estimators $\hat{c}_{CME}$ and $\hat{\lambda}_{CME}$ of the parameters c and λ are obtained by minimizing, with respect to c and λ, the function:

$$C(c,\lambda) = \frac{1}{12n} + \sum_{i=1}^{n}\left(F\left(x_{i:n}\,|\,c,\lambda\right) - \frac{2i-1}{2n}\right)^2. \tag{38}$$

These estimators can also be obtained by solving the following non-linear equations:

$$\sum_{i=1}^{n}\left(F\left(x_{i:n}\,|\,c,\lambda\right) - \frac{2i-1}{2n}\right)\eta_1\left(x_{i:n}\,|\,c,\lambda\right) = 0,$$

$$\sum_{i=1}^{n}\left(F\left(x_{i:n}\,|\,c,\lambda\right) - \frac{2i-1}{2n}\right)\eta_2\left(x_{i:n}\,|\,c,\lambda\right) = 0,$$

where $\eta_1\left(\cdot\,|\,c,\lambda\right)$ and $\eta_2\left(\cdot\,|\,c,\lambda\right)$ are given by (30) and (31), respectively.

3.8.2 Methods of Anderson-Darling and Right-tail Anderson-Darling

Other type of maximum goodness-of-fit estimator is based on an Anderson-Darling statistic (Anderson & Darling, (1952)) and is known as the Anderson-Darling estimator. The Anderson-Darling estimators $\hat{c}_{ADE}$ and $\hat{\lambda}_{ADE}$ of the parameters c and λ are obtained by minimizing, with respect to c and λ, the function:

$$\begin{aligned} A(c,\lambda) = & -n - \frac{1}{n}\sum_{i=1}^{n}(2i-1)\Big\{\log F\left(x_{i:n}\,|\,c,\lambda\right) + \\ & \log \overline{F}\left(x_{n+1-i:n}\,|\,c,\lambda\right)\Big\}. \end{aligned} \tag{39}$$

These estimators can also be obtained by solving the following non-linear equations:

$$\sum_{i=1}^{n}(2i-1)\left[\frac{\eta_1\left(x_{i:n}\,|\,c,\lambda\right)}{F\left(x_{i:n}\,|\,c,\lambda\right)} - \frac{\eta_1\left(x_{n+1-i:n}\,|\,c,\lambda\right)}{\overline{F}\left(x_{n+1-i:n}\,|\,c,\lambda\right)}\right] = 0,$$

$$\sum_{i=1}^{n}(2i-1)\left[\frac{\eta_2\left(x_{i:n}\,|\,c,\lambda\right)}{F\left(x_{i:n}\,|\,c,\lambda\right)} - \frac{\eta_2\left(x_{n+1-i:n}\,|\,c,\lambda\right)}{\overline{F}\left(x_{n+1-i:n}\,|\,c,\lambda\right)}\right] = 0,$$

where $\eta_1(\cdot|c,\lambda)$ and $\eta_2(\cdot|c,\lambda)$ are given by (28) and (29) , respectively.

Similarly, the Right-tail Anderson-Darling estimators $\hat{c}_{RTADE}$ and $\hat{\lambda}_{RTADE}$ of the parameters c and λ can be obtained by minimizing, with respect to c and λ, the function:

$$R(c,\lambda) = \frac{n}{2} - 2\sum_{i=1}^{n} F\left(x_{i:n} \,|\, c,\lambda\right) - \frac{1}{n}\sum_{i=1}^{n}(2i-1)$$
$$\log \overline{F}\left(x_{n+1-i:n} \,|\, c,\lambda\right). \tag{40}$$

These estimators can be obtained by solving the following non-linear equations:

$$-2\sum_{i=1}^{n}\frac{\eta_1\left(x_{i:n}\,|\,c,\lambda\right)}{F\left(x_{i:n}\,|\,c,\lambda\right)} + 1n\sum_{i=1}^{n}(2i-1)\frac{\eta_1\left(x_{n+1-i:n}\,|\,c,\lambda\right)}{\overline{F}\left(x_{n+1-i:n}\,|\,c,\lambda\right)} = 0,$$

$$-2\sum_{i=1}^{n}\frac{\eta_2\left(x_{i:n}\,|\,c,\lambda\right)}{F\left(x_{i:n}\,|\,c,\lambda\right)} + 1n\sum_{i=1}^{n}(2i-1)\frac{\eta_2\left(x_{n+1-i:n}\,|\,c,\lambda\right)}{\overline{F}\left(x_{n+1-i:n}\,|\,c,\lambda\right)} = 0,$$

where $\eta_1(\cdot|c,\lambda)$ and $\eta_2(\cdot|c,\lambda)$ are given by (28) and (29), respectively.

4. Bayesian Analysis

In this section, we consider Bayesian inference of the unknown parameters of the $GM(c, \lambda)$. It is assumed that c and λ has the independent gamma prior distributions with probability density functions

$$g(c) \propto c^{\alpha-1}e^{-c\beta} \qquad\qquad c > 0 \tag{41}$$

and

$$g(\lambda) \propto \lambda^{a-1}e^{-b\lambda} \qquad\qquad \lambda > 0. \tag{42}$$

The hyperparameters a, b, α, and β are known and non-negative. If both parameters c and λ are unknown, joint conjugate priors do not exist. It is not unreasonable to assume independent gamma priors on the shape and scale parameters for a two-parameter $GM(c, \lambda)$, because gamma distributions are very flexible, and the Jeffreys (non-informative) prior, introduced by Jeffreys (1946) is a special case of this. The joint prior distribution for both parameters in this case is given by

$$p(c,\lambda) \propto c^{\alpha-1}\exp(-c\beta)\lambda^{a-1}\exp(-b\lambda).$$

Thus, the joint posterior distribution is given by

$$p(c, \lambda \mid \mathbf{x}) \ \propto \ \lambda^{n+a-1} c^{\alpha-1} \, e^{c\left(\sum_{i=1}^{n} x_i - \beta\right)} e^{-\lambda\left[\frac{1}{c}\sum_{i=1}^{n}(e^{cx_i}-1)+b\right]}. \tag{43}$$

The marginal distribution of c given λ and data is given by

$$p(c \mid \lambda, \mathbf{x}) \ \propto \ c^{\alpha-1} \, e^{c\left(\sum_{i=1}^{n} x_i - \beta\right)} e^{-\frac{\lambda}{c}\sum_{i=1}^{n}(e^{cx_i}-1)}.$$

Similarly, the marginal distribution of λ given c and data is given by

$$p(\lambda \mid c, \mathbf{x}) \ \propto \ \lambda^{n+a-1} e^{-\lambda\left[\frac{1}{c}\sum_{i=1}^{n}(e^{cx_i}-1)+b\right]}.$$

Next, we must consider the question of what loss function will be used to derive the Bayes estimates from the marginal posterior distributions. Below we provide an MCMC algorithm and the loss functions (See Ali et al. (2013); Dey et al. (2015)) used to compute the Bayes estimates and their respective posterior risk (see Table 2).

4.1 *Algorithm*

The algorithm used to obtain the Bayesian estimates is given as follows:

- Set initial values c_0 and λ_0 for the parameters c and λ;
- At step $i + 1$, we draw a new set of values $(c_{i+1},\ \lambda_{i+1})$ from the conditional posterior $p(c_i \mid \lambda_i, \mathbf{x})$ and $p(\lambda_i \mid c_i, \mathbf{x})$;
- Repeat above step M times to obtain one chain of M values for (c, λ);
- The Bayes estimate of the parameters c and λ, under the above-defined loss functions are given as

$$\hat{c} \ = \ \frac{1}{M}\sum_{i=1}^{M} c_i \, , \ \hat{\lambda} = \frac{1}{M}\sum_{i=1}^{M} \lambda_i \, , \tag{44}$$

   respectively.
- The credible interval can be obtained by sorting the observations of c and λ.

5. Simulations

In this section we present some experimental results to evaluate the performance of the different methods of estimation discussed in the

previous sections. The study is based on 1000 generated data set from the $GM(c, \lambda)$ for different sample sizes (n = 10, 50, 100) and different values of (c, λ). The values of true parameters are chosen to propose the different shape of density function. The estimates are compared in terms of their bias and mean-squared error (MSE). Tables 3-4 contains the averages of the 1000 bias and the corresponding mean squared error (MSE).

We also obtained the Bayes estimates under five different loss functions such as, the squared error loss function (SELF), weighted squared error loss function (WSELF), modified squared error loss function (MSELF), precautionary loss function (PLF) and K-loss (KLF).

Estimates and respective posterior risks under different loss functions are calculated for the simulated data set and are presented in Tables 7-8. Figures 3-4 show the risk functions of the estimates for different sample sizes and for different values of the parameters c and λ.

To obtain the Bayes estimates and credible intervals, we need to appeal to the MCMC algorithm in order to obtain a sample of values of c and λ from the joint posterior. Specifically, a chain is run for 20,000 iterations and the first 5,000 runs are discarded, and base inferences on posterior summaries of the parameters and credible intervals calculated from the samples.

Besides, the empirical coverage probability for parameters c and λ are obtained. Tables 5-6 show the coverage probabilities varying the sample size n for the frequentist methods discussed in the paper. The confidence intervals for the parameters are constructed by using asymptotic distribution and also by using the bootstrap approach. Some of the points are quite clear from Tables 3-4 that as sample size increases, the average bias and the MSEs decrease. It verifies the consistency of all the estimators.

The bias and MSEs of c generally increases with increasing c and λ for all methods of estimation. The bias and MSEs of λ generally increases with increasing λ and c for all methods of estimation.

When c increases and λ decreases, MMM estimators have the smallest MSEs for c and MME estimators have the smallest MSEs for λ. The Performance of the MPS are also quite satisfactory.

It is evident from Tables 7-8 that as c increases and λ decreases, KLF has the smallest posterior risk as compare to other loss functions while PLF is on the second choice for c but for λ, MSELF has the smallest posterior risk as compare to other loss functions while KLF is the second choice.

Also, when both c and λ increases, KLF has the smallest risk in case of c and MSELF has the smallest risk in case of λ. It is to be noted that

posterior risk is more comprehensive measure for comparison of different loss functions in Bayesian setup.

6. An example with literature data

In this section, we use a real data set to illustrate the proposed estimation methods discussed in the previous sections. Let us consider the following data set introduced in King et al. (1979):

112, 68, 84, 109, 153, 143, 60, 70, 98, 164, 63, 63, 77, 91, 91, 66, 70, 77, 63, 66, 66, 94, 101, 105, 108, 112, 115, 126, 161, 178.

These data represent the numbers of tumor-days of 30 rats fed with unsaturated diet. Chen (1997) and Asgharzadeh and Abdi (2011) used the Gompertz distribution for these data set in order to obtain exact confidence intervals and joint confidence regions for the parameters based on two different statistical analysis. Let us also assume the Gompertz distribution with density (1) fitted to the data and to compare the performance of the methods discussed in this paper. For a Bayesian analysis, we assume the independent Gamma prior distributions for the parameters c and λ, with the hyper parameter values $a = b = \alpha = \beta = 0.01$. The Bayes estimates cannot be obtained in closed form therefore we use MCMC procedure to compute the Bayes estimates and also to construct credible intervals. Using the software R, we simulated 50,000 MCMC samples (5.000 "burn-in-samples") for the joint posterior distribution. The convergence of the chains was monitored from trace plots of the simulated samples. The estimates and 95% confidence intervals under classical methods are given in Table 9. We also computed the Bayesian estimates and 95% intervals credible intervals for the parameters based on MCMC algorithm. The estimated values are $\hat{c} = 0.0239$, $\hat{\lambda} = 0.0017$ and the corresponding intervals are given by (0.0154, 0.0320) and (0.0006, 0.0038). The marginal posterior distributions for the parameters c and λ of the Gompertz distribution are shown in Figure 5. Table 10 shows the Bayes estimators and corresponding risks obtained with respect to symmetric and asymmetric loss functions.

Real data analysis shows that MME performs better than among all frequentist estimators while MMM is the second choice in case of c. But in case of λ, WLSE performs better than among all the frequentist estimators while LME is the second choice. It is noticed from the real data analysis that both MSELF and KLF have the smallest risk among all Bayes estimators considered.

7. Conclusion

In this article, we provide explicit expressions for the quantiles, moments, moment generating function, hazard rate, mean residual lifetime, mean past lifetime, stochastic ordering, various entropies, stress-strength parameter, Bonferroni and Lorenz curve and order statistics. The model parameters are estimated by fifteen methods of estimation, namely maximum likelihood estimators, moments estimators, pseudo-moments estimators, modified moments estimators, L-moment estimators, percentile based estimators, least squares and weighted least squares estimators, maximum product of spacings estimators, minimum spacing absolute distance estimators, minimum spacing absolute-log distance estimator, Cramér-von-Mises estimators, Anderson-Darling and right-tail Anderson-Darling estimators and the Bayes estimators. Results of a simulation study to compare these methods are presented, which showed that among frequentist estimators, modified moments and moments estimators performs better than their counter parts. Among the Bayes estimators, Bayes estimators based on MSELF and KLF performed better. The performance of maximum product spacing estimators is fairly reasonable and competitive.

Acknowledgment
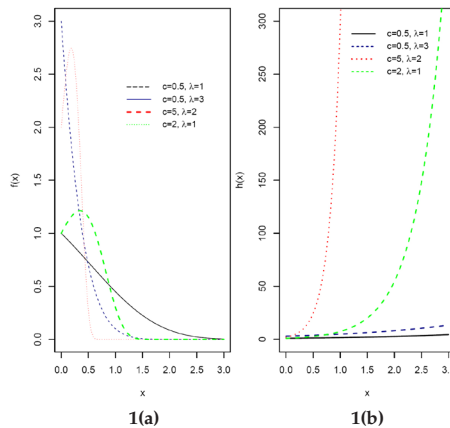
Figure 1
1(a):pdf and 1 (b):hazard function ($h(x)$)
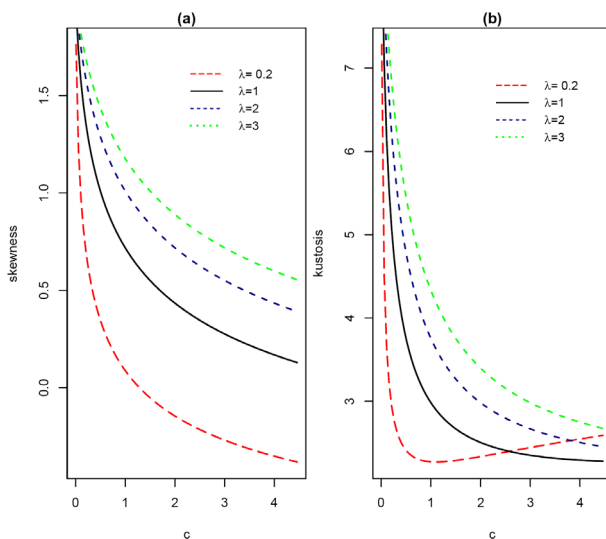
Figure 2

Plots of skewness and kurtosis versus c
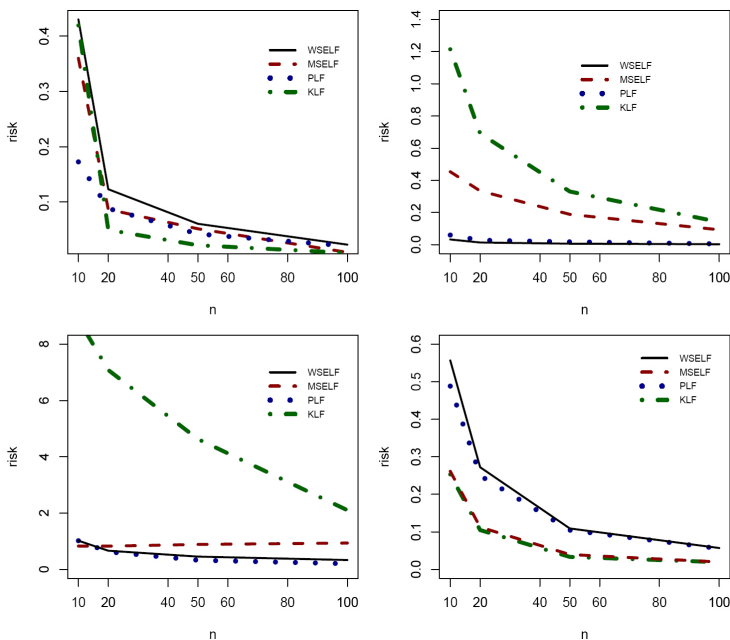
Figure 3

Plot of risk functions for different sample sizes for the parameters c and λ.
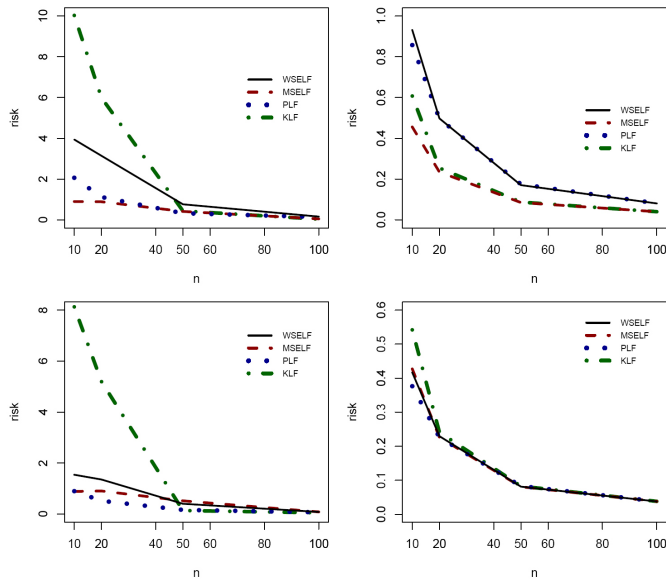
Figure 4

Plot of risk functions for different sample sizes for the parameters c and λ.
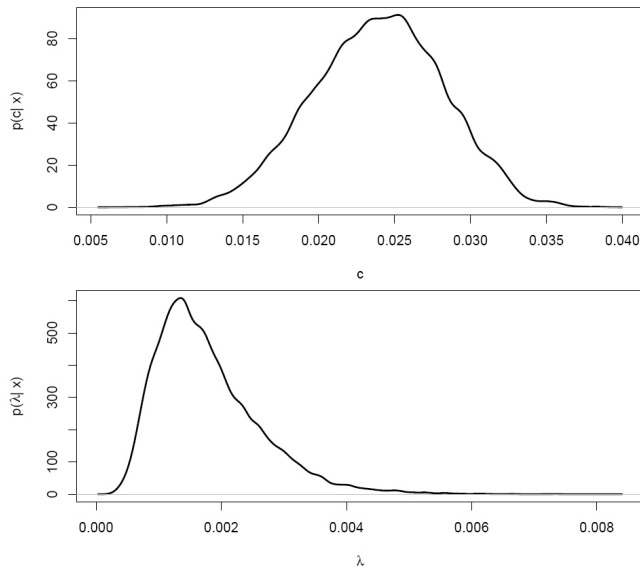
Figure 5

The posterior densities for the parameters c and λ of the Gompertz
distribution fitted by the data.

Appendix

Table 1

Mean, variance, skewness, kurtosis and coefficient of variation (CV) for different values of λ and $c = 5$

λ	Mean	Variance	Skewness	Kurtosis	CV
0.1	0.040483	0.119325	137.5606	163.1597	8.532829
0.2	0.081943	0.237156	34.09007	79.87667	5.942993
0.5	0.212367	0.579386	6.266150	30.03675	3.584239
1.0	0.450999	1.017066	6.094266	13.68068	2.321868
1.5	0.718330	1.075730	5.345578	8.776922	1.712642
2.0	1.016999	1.096544	1.331925	8.520371	1.312654
2.5	1.349859	1.513496	0.930739	7.116877	1.005465
3.0	1.719995	1.618114	0.632212	6.842248	0.739567
3.5	2.130746	1.782139	0.464134	5.520244	0.473307
4.0	2.585719	1.842089	0.438354	5.40.998	0.401116

Table 2

Bayes estimator and posterior risk under different loss functions

Loss function	Bayes estimator	Posterior risk
$L_1 = SELF = (\theta - d)^2$	$E(\theta\|\mathbf{x})$	$\text{var}(\theta\|\mathbf{x})$
$L_2 = WSELF = \dfrac{(\theta - d)^2}{\theta}$	$(E(\theta^{-1}\|\mathbf{x}))^{-1}$	$E(\theta\|\mathbf{x}) - (E(\theta^{-1}\|\mathbf{x}))^{-1}$
$L_3 = MSELF = \left(1 - \dfrac{d}{\theta}\right)^2$	$\dfrac{E(\theta^{-1}\|\mathbf{x})}{E(\theta^{-2}\|\mathbf{x})}$	$1 - \dfrac{E(\theta^{-1}\|\mathbf{x})^2}{E(\theta^{-2}\|\mathbf{x})}$
$L_4 = PLF = \dfrac{(\theta - d)^2}{d}$	$\sqrt{E(\theta^2\|\mathbf{x})}$	$2(\sqrt{E(\theta^2\|\mathbf{x})} - E(\theta\|\mathbf{x}))$
$L_5 = KLF = \left(\sqrt{\dfrac{d}{\theta}} - \sqrt{\dfrac{\theta}{d}}\right)^2$	$\sqrt{\dfrac{E(\theta\|\mathbf{x})}{E(\theta^{-1}\|\mathbf{x})}}$	$2(\sqrt{E(\theta\|\mathbf{x})E(\theta^{-1}\|\mathbf{x})} - 1)$

Note: SELF, squared error loss function; WSELF, weighted squared error loss function; MSELF, modified squared error loss function; PLF, precautionary loss function, and KLF, K-loss function.

Table 3

Average bias of the estimates of c and λ and their associated MSEs (in parenthesis) for the different methods with $c = 3$ and $\lambda = 0.02$

Method	$c = 3$			$\underline{\lambda} = 0.02$		
	$n = 10$	$n = 50$	$n = 100$	$n = 10$	$n = 50$	$n = 100$
MLE	0.4853	0.1003	0.0532	0.0040	0.0004	-9.6e-05
	(1.4217)	(0.1556)	(0.0685)	(0.0009)	(0.0001)	(6.0e-05)
MME	0.1117	0.0622	0.0627	0.0084	0.0019	7.9e-05
	(0.4145)	(0.1741)	(0.0958)	(0.0014)	(0.0002)	(8.0e-05)
PME	0.4458	0.1020	0.0699	0.0068	0.0016	1.2e-05
	(1.2453)	(0.2382)	(0.1034)	(0.0015)	(0.0002)	(8.1e-05)
MMM	-0.0578	0.0039	0.0038	0.0147	0.0038	0.0019
	(0.8231)	(0.2305)	(0.1258)	(0.0021)	(0.0002)	(0.0001)
LSE	-0.0376	-0.0033	0.0104	0.0238	0.0041	0.0014
	(1.3166)	(0.2133)	(0.0956)	(0.0031)	(0.0002)	(8.9e-05)
WLSE	-0.0042	0.0253	0.0269	0.0215	0.0028	0.0007
	(1.2806)	(0.1788)	(0.0796)	(0.0027)	(0.0002)	(7.2e-05)
PCE	-0.1969	-0.0857	-0.0376	0.0274	0.0063	0.0025
	(0.9791)	(0.1854)	(0.0819)	(0.0032)	(0.0002)	(9.1e-05)
LME	-0.0665	-0.0590	-0.0413	0.0219	0.0060	0.0032
	(0.6333)	(0.1993)	(0.1144)	(0.0036)	(0.0004)	(0.0001)
MPS	-0.2515	-0.1042	-0.0661	0.0270	0.0060	0.0030
	(0.8423)	(0.1392)	(0.0656)	(0.0030)	(0.0002)	(8.4e-05)
MSADE	-0.2350	-0.1118	-0.0605	0.0334	0.0080	0.0039
	(1.0084)	(0.2155)	(0.1154)	(0.0056)	(0.0004)	(0.0001)
MSALDE	-0.2222	-0.0919	-0.0611	0.0321	0.0068	0.0035
	(0.9501)	(0.1815)	(0.0884)	(0.0051)	(0.0003)	(0.0001)
CME	0.4990	0.0933	0.0581	0.0088	0.0016	0.0003
	(2.0563)	(0.2368)	(0.1020)	(0.0016)	(0.0002)	(8.1e-05)
ADE	0.0659	0.0296	0.0272	0.0673	0.0027	0.0007
	(5.9452)	(0.1719)	(0.0757)	(2.8688)	(0.0002)	(7.1e-05)
RTADE	0.2880	0.0550	0.0316	0.0129	0.0020	0.0007
	(1.5559)	(0.1760)	(0.0830)	(0.0024)	(0.0002)	(8.1e-05)

Table 4

Average bias of the estimates of c and λ and their associated MSEs (in parenthesis) for the different methods with $c = 0.5$ and $\lambda = 3$

Method	$c = 0.5$			$\underline{\lambda} = 3$		
	$n = 10$	$n = 50$	$n = 100$	$n = 10$	$n = 50$	$n = 100$
MLE	1.3604	0.2068	0.0678	-0.3588	-0.1174	-0.0503
	(7.1912)	(0.4408)	(0.1994)	(2.1338)	(0.3379)	(0.1763)
MME	1.6132	0.4784	0.3307	-0.6285	-0.31423	-0.2339
	(6.7009)	(0.5393)	(0.3065)	(1.4136)	(0.3642)	(0.2220)
PME	1.0469	0.4574	0.3364	-0.3165	-0.3102	-0.2163
	(2.0004)	(0.3564)	(0.2237)	(1.1742)	(0.3000)	(0.1795)
MMM	1.0793	0.2958	0.1711	0.8509	0.4362	0.3967
	(4.1126)	(0.3251)	(0.1366)	(11.3500)	(1.8746)	(1.2864)
LSE	-0.0533	-0.0526	-0.0250	0.5580	0.1048	0.0644
	(10.4377)	(1.0366)	(0.4675)	(5.1551)	(0.6531)	(0.3116)
WLSE	0.0769	0.0345	0.0196	0.4802	0.0585	0.0374
	(9.2030)	(0.7560)	(0.3264)	(4.6047)	(0.5627)	(0.2750)
PCE	-0.0190	-0.1296	-0.1241	0.3130	0.1247	0.1177
	(4.7496)	(0.4337)	(0.2125)	(3.3975)	(0.5131)	(0.2686)
LME	1.3415	0.6158	0.4276	-0.5846	-0.4392	-0.3260
	(5.4819)	(0.7865)	(0.4624)	(1.1810)	(0.4475)	(0.2852)
MPS	-0.1630	-0.1595	-0.1337	0.4382	0.1357	0.1123
	(4.4940)	(0.3987)	(0.1882)	(3.5166)	(0.4526)	(0.2326)
MSADE	0.4820	0.0055	-0.0241	0.4800	0.1073	0.0832
	(7.9963)	(0.6330)	(0.3013)	(6.1932)	(0.7594)	(0.3763)
MSALDE	-0.1156	-0.1564	-0.1245	0.5536	0.1753	0.1263
	(5.5052)	(0.5122)	(0.2395)	(4.7821)	(0.6586)	(0.3321)
CME	1.4440	0.1624	0.0812	0.0372	0.0175	0.0210
	(18.2589)	(1.1202)	(0.4873)	(4.0886)	(0.6287)	(0.3045)
ADE	0.2725	0.0144	0.0008	0.3821	0.0738	0.0499
	(6.5831)	(0.6162)	(0.2749)	(4.7571)	(0.5467)	(0.2681)
RTADE	0.5669	0.0655	0.0095	0.1764	0.0257	0.0404
	(6.3753)	(0.4781)	(0.2251)	(3.6210)	(0.4624)	(0.2444)

Table 5
Coverage probabilities for the parameters c and λ

	$c = 3$			$\underline{\lambda} = 0.02$		
Method	$n = 10$	$n = 50$	$n = 100$	$n = 10$	$n = 50$	$n = 100$
MLE	0.95	0.95	0.95	0.73	0.87	0.91
MME	1.00	0.98	0.97	1.00	0.97	0.95
PME	0.92	0.93	0.94	0.91	0.93	0.93
MMM	0.93	0.96	0.96	0.94	0.96	0.96
LSE	0.95	0.93	0.95	0.95	0.93	0.95
WLSE	0.95	0.94	0.95	0.94	0.94	0.95
PCE	0.94	0.91	0.95	0.93	0.92	0.95
LME	0.94	0.94	0.94	0.95	0.93	0.95
MPS	0.91	0.91	0.91	0.92	0.91	0.93
MSADE	0.97	0.91	0.93	0.95	0.92	0.94
MSALDE	0.93	0.90	0.92	0.92	0.91	0.93
CME	0.86	0.91	0.94	0.87	0.93	0.94
ADE	0.94	0.94	0.94	0.93	0.94	0.95
RTADE	0.90	0.94	0.94	0.91	0.94	0.94

Table 6
Coverage probabilities for the parameters c and λ

	$c = 0.5$			$\underline{\lambda} = 3$		
Method	$n = 10$	$n = 50$	$n = 100$	$n = 10$	$n = 50$	$n = 100$
MLE	0.92	0.95	0.95	0.85	0.93	0.95
MME	0.73	0.81	0.80	0.90	0.93	0.90
PME	0.90	0.99	0.92	0.95	0.97	0.93
MMM	0.81	0.92	0.95	0.89	0.95	0.96
LSE	0.94	0.96	0.98	0.96	0.97	0.98
WLSE	0.91	0.80	0.93	0.95	0.82	0.94
PCE	0.78	0.87	0.94	0.79	0.85	0.93
LME	0.92	0.96	0.97	0.95	0.96	0.97
MPS	0.63	0.89	0.92	0.63	0.87	0.92
MSADE	0.67	0.87	0.93	0.69	0.87	0.92
MSALDE	0.57	0.73	0.83	0.58	0.72	0.81
CME	0.56	0.70	0.83	0.59	0.72	0.85
ADE	0.62	0.74	0.86	0.63	0.75	0.86
RTADE	0.57	0.73	0.85	0.60	0.74	0.86

Table 7

Bayes estimates of c and λ and their posterior risks under different loss
functions for $c = 3$ and $\lambda = 0.02$

		$c = 3$			$\underline{\lambda} = 0.02$		
	Method	$n = 10$	$n = 50$	$n = 100$	$n = 10$	$n = 50$	$n = 100$
	SELF	3.3000	3.0755	3.0403	0.0458	0.0238	0.0215
	WSELF	2.9325	3.0317	3.0193	0.0138	0.0180	0.0186
Estimate	MSELF	2.5792	2.9871	2.9981	0.0041	0.0133	0.0160
	PLF	3.4207	3.0970	3.0507	0.0700	0.0271	0.0230
	KLF	3.0807	3.0536	3.0298	0.0245	0.0207	0.0200
	Method	$n = 10$	$n = 50$	$n = 100$	$n = 10$	$n = 50$	$n = 100$
	SELF	0.8214	0.1333	0.0634	0.0045	0.0001	7.3e-05
	WSELF	0.3674	0.0438	0.0209	0.0320	0.0058	0.0028
Risk	MSELF	0.1836	0.0148	0.0070	0.7583	0.2853	0.1453
	PLF	0.2414	0.0428	0.0207	0.0482	0.0065	0.0030
	KLF	0.7736	0.0145	0.0069	3.3233	0.3452	0.1572

Table 8

Bayes estimates of c and λ and their posterior risks under different loss
functions for $c = 0.5$ and $\lambda = 3$

		$c = 0.5$			$\underline{\lambda} = 3$		
	Method	$n = 10$	$n = 50$	$n = 100$	$n = 10$	$n = 50$	$n = 100$
	SELF	1.1078	0.6283	0.5556	3.1018	2.9701	2.9884
	WSELF	0.0430	0.1750	0.2618	2.5418	2.8587	2.9334
Estimate	MSELF	0.0034	0.0159	0.0279	1.9664	2.7395	2.8778
	PLF	1.6181	0.7756	0.6419	3.3470	3.0230	3.0154
	KLF	0.1952	0.3249	0.3782	2.8030	2.9138	2.9608
	Method	$n = 10$	$n = 50$	$n = 100$	$n = 10$	$n = 50$	$n = 100$
	SELF	2.1225	0.2275	0.1101	1.6643	0.3158	0.1613
	WSELF	1.0648	0.4532	0.2938	0.5599	0.1114	0.0550
Risk	MSELF	0.8538	0.8987	0.9142	0.2647	0.0428	0.0193
	PLF	1.020	0.2945	0.1724	0.4905	0.1056	0.0540
	KLF	8.9484	2.2521	1.1342	0.2571	0.0399	0.0190

Table 9
Estimators and 95% confidence intervals of c and λ of the Gompertz distribution for different estimation methods

Method	$\hat{c}$	95% CI	94λ	95% CI
MLE	0.0241	(0.0160, 0.0322)	0.0016	(0.0002, 0.0031)
MME	0.0113	(0.0097, 0.0125)	0.0048	(0.0034, 0.0067)
PME	0.0116	(0.0102, 0.0119)	0.0047	(0.0038, 0.0073)
MMM	0.0134	(0.0090, 0.0137)	0.0034	(0.0026, 0.0061)
LSE	0.0332	(0.0210, 0.0501)	0.0010	(0.0002,0.0030)
WLSE	0.0416	(0.0287, 0.0612)	0.0005	(0.0001, 0.0016)
PCE	0.0291	(0.0187, 0.0403)	0.0011	(0.0003, 0.0030)
LME	0.0329	(0.0202, 0.0456)	0.0008	(0.0002, 0.0028)
MPS	0.0214	(0.0126, 0.0298)	0.0019	(0.0007, 0.0043)
MSADE	0.0215	(0.0116, 0.0318)	0.0016	(0.0005, 0.0049)
MSALDE	0.0190	(0.0099, 0.0285)	0.0026	(0.0009, 0.0066)
CME	0.0355	(0.0242, 0.0538)	0.0009	(0.0002, 0.0025)
ADE	0.0243	(0.0157, 0.0367)	0.0017	(0.0005, 0.0040)
RTADE	0.0208	(0.0130, 0.0327)	0.0023	(0.0006, 0.0049)

Table 10
Bayes estimates of c and λ and posterior risks under different loss functions

Loss functions	$\hat{c}$	risk	$\hat{\lambda}$	risk
SELF	0.0239	0.0042	0.0018	0.0008
WSELF	0.0230	0.0008	0.0014	0.0003
MSELF	0.0220	0.0404	0.0011	0.1961
PLF	0.0242	0.0008	0.0020	0.0004
KLF	0.0234	0.0363	0.0016	0.2203

References

[1] AL-Eideh, B. M. (1996). Asymptotic normality of maximum likelihood estimators for non-parametric Markov chains, *Journal of Information and Optimization Sciences*, 17(2), 233-238.

[2] Ali, S., Aslam, M. and Ali Kazmi, S. M. (2013). A study of the effect of the loss function on Bayes Estimate, posterior risk and hazard

function for Lindley distribution, *Applied Mathematical Modelling*, 37, 6068–6078.

[3] Alkasasbeh, M. R. and Raqab, M. Z.(2009). Estimation of the generalized logistic distribution parameters: comparative study, *Statistical Methodology*; 6, 262-279.

[4] Anderson, T. W. and Darling, D. A. (1952). Asymptotic theory of certain "goodness-of-fit" criteria based on stochastic processes, *Annals of Mathematical Statistics*, 23, 193-212.

[5] Asgharzadeh, A and Abdi, M. (2011). Exact Confidence Intervals and Joint Confidence Regions for the Parameters of the Gompertz Distribution based on Records. *Pak. J. Statist.*, 27(1), 55-64.

[6] Berkson J. (1944). Application of the logistic function to bioassay. *Journal of the American Statistical Association*; 39, 357-365.

[7] Bonferroni C.E. (1930). Elmenti di statistica generale. Libreria Seber, Firenze.

[8] Casella, G. and Berger, R. L. (1990). Statistical Inference. *Brooks/Cole Publishing Company*, California.

[9] Chen, Z. (1997). Parameter estimation of the Gompertz population, *Biometrical Journal*. 39, 117-124.

[10] Cheng, R. C. H. and Amin, N. A. K. (1979). Maximum product-of-spacings estimation with applications to the lognormal distribution, *Technical Report, Department of Mathematics, University of Wales*.

[11] Cheng, R. C. H. and Amin, N. A. K. (1983). Estimating parameters in continuous univariate distributions with a shifted origin, *Journal of the Royal Statistical Society*, Ser B, 3, 394-403.

[12] Dey, S., Dey, T. and Kundu D. (2014). Two-parameter Rayleigh distribution: Different methods of estimation. Amer J Math Manag Sci. 33, 55-74.

[13] Dey, S., Ali, S. and Park, C. (2015) Weighted exponential distribution: properties and different methods of estimation. J Stat Comput Simul, 85, 3641-3661.

[14] Dey, S., Dey. T., Ali, S. and Mulekar, M. S. (2016). Two-parameter Maxwell distribution: Properties and different methods of estimation. J Stat Theory Prac, 10, 291-310.

[15] Dey, S., Kumar, D., Ramos, P. L. and Louzada, F. (2017a). Exponentiated Chen distribution: Properties and estimation. Commun. Stat. Simul. Comput., 46, 8118-8139.

[16] Dey, S., Al-Zahrani, B. and Basloom, S. (2017b). Dagum distribution: Properties and different methods of estimation. Int J Stat Probab, 6, 74-92.

[17] Dey, S., Raheem, E. and Mukherjee, S. (2017c). Statistical properties and different methods of estimation of transmuted Rayleigh distribution. Rev. Colombiana Estadíst. 40, 165-203.

[18] Dey, S., Raheem, E., Mukherjee, S. and Ng, H.K.T. (2017d). Two parameter exponentiated-Gumbel distribution: properties and estimation with flood data example. J Statist Manag Sys, 20, 197-233.

[19] Dey, S., Alzaatreh, A., Zhang, C. and Kumar, D. (2017e). A New Extension of Generalized Exponential Distribution with Application to Ozone Data. Ozone: Science & Engineering, 39, 273-285.

[20] Dey, S., Zhang, C., Asgharzadeh, A. and M. Ghorbannezhad, M. (2017f). Comparisons of Methods of Estimation for the NH Distribution. Annals of Data Science (to appear).

[21] DÁgostino, R.B. and Stephens, M.A. (1986). Goodness-of-fit Techniques, *Marcel Dekker, New York*.

[22] Johnson, R.C. and Johnson, N.L. (1979). Survival Models and Data Analysis, John Wiley, New York.

[23] Franses, P.H.(1994). Fitting a Gompertz curve, *Journal of the Operational Research Society*, 45, 109-113.

[24] Garg, M., Rao, B. and Redmond, C. (1970). Maximum-likelihood estimation of the parameters of the Gompertz survival function. *Journal of the Royal Statistical Society*, Series C (Applied Statistics) 19 (2), 152-159.

[25] Gompertz, B. (1825). On the nature of the function expressive of the law of human mortality, and on a new mode of determining the value of life contingencies. *Philosophical Transactions of the Royal Society of London*, 115, 513-583.

[26] Hosking, J. R. M. (1990). L-moments: Analysis and estimation for distributions using luiear combinations of order statistics, *Journal of the Royal Statistical Society*, Ser, B 52, 105-124.

[27] Ismail, A. A. (2010). Bayes estimation of Gompertz distribution parameters and acceleration factor under partially accelerated life tests with type-I censoring. *Journal of Statistical Computation and Simulation*, 80(11), 1253-1264.

[28] Jaheen, Z.F. (2003a). Prediction of Progressive Censored Data from the Gompertz Model. *Communications in Statistics Simulation and Computation*, 32(3), 663-676.

[29] Jaheen, Z.F. (2003b). A Bayesian analysis of record statistics from the Gompertz model. *Applied Mathematics and Computation*, 145, 307-320.

[30] Jeffreys, H.(1946). An invariant form for the prior probability in estimation problems. *Proc R Soc London. Ser A Math Phy Sci.* ; 186(1007):435-461.

[31] Kao, J.H.K. (1958). Computer methods for estimating Weibull parameters in reliability studies, *Trans. IRE Reliability Quality Control*, 13, 15-22.

[32] Kao, J.H.K. (1959). A graphical estimation of mixed Weibull parameters in life testing electron tube, *Technometrics*, 1, 389-407.

[33] Kiani, K., Arasan, J., Midi, H. (2012). Interval estimations for parameters of Gompertz model with time-dependent covariate and right censored data. *Sains Malaysiana*, 41(4), 471-480.

[34] King, M., Bailey, D.M., Gibson, D.G., Pitha, J.V. and McCay, P.B. (1979). Incidence and growth of mammary tumors induced by 7, l2-Dimethylbenz antheacene as related to the dietary content of fat and antioxidant. *Journal of National Cancer Institute*, 63, 656-664.

[35] Kundu, D. and Raqab, M.Z. (2005). Generalized Rayleigh distribution: Different methods of estimation, *Computational Statistics and Data Analysis*, 49, 187-200.

[36] Lenart, A. (2014). The moments of the Gompertz distribution and maximum likelihood estimation of its parameters. *Scandinavian Actuarial Journal*, 3, 255-277.

[37] Luceño, A. (2006). Fitting the Generalized Pareto Distribution to Data using Maximum Goodness-of-fit Estimators. *Computational Statistics and Data Analysis*, 51(2), 904-917.

[38] Mann, N. R., Schafer, R. E. and Singpurwalla, N. D. (1974). Methods for Statistical Analysis of Reliability and Life Data, *New York, Wiley*.

[39] Makany, R. A. (1991). Theoretical basis of Gompertz's curve, *Biometrical Journal*. 33, 121-128.

[40] Milgram, M. (1985). The generalized integro-exponential function. *Mathematics of Computation*, 44 (170), 443-458.

[41] MacDonald, P. D. M. (1971). Comment on "An estimation procedure for mixtures of distributions" by Choi and Bulgren, *Journal of the Royal Statistical Society B*, 33, 326-329.

[42] Mufid, M. Azzam. (1994). Maximum Likelihood Estimates Using Different Experimental Schemes, *Journal of Information and Optimization Sciences*, 15(3), ,431-442.

[43] Ojo M. O.(1997). Some relationships between the generalized logistic and other distributions. *Statistica*, LVII: 573-579.

[44] Ranneby, B. (1984). The maximum spacing method. An estimation method related to the maximum likelihood method, *Scandinavian Journal of Statistics*, 11, 93-112.

[45] Rényi, A. (1961). On measures of entropy and information. *In: Proceedings of the 4th Berkeley Symposium on Mathematical Statistics and Probability*, I, 547-561. University of California Press, Berkeley.

[46] Rao, B.R. and C.V. Damaraju, C.V.(1992). New better than used and other concepts for a class of life distribution, *Biometrical Journal*, 34 , 919-935.

[47] Read, C.B. (1983). Gompertz Distribution, Encyclopedia of Statistical Sciences, *Wiley, New York*.

[48] Shaked M, Shanthikumar J.G.(1994). Stochastic orders and their applications. *Boston, MA: Academic Press*.

[49] Shannon, C. E. (1951). Prediction and entropy of printed English. *The Bell System Technical Journal*, 30, 50-64.

[50] Swain, J., Venkatraman, S. and Wilson, J. (1988). Least squares estimation of distribution function in Johnson's translation system, *Journal of Statistical Computation and Simulation*, 29, 271-297.

[51] Teimouri, M., Hoseini, S. M. and Nadarajah, S. (2013). Comparison of estimation methods for the Weibull distribution, *Statistics*, 47, 93-109.

[52] Torabi, H. (2008). A general method for estimating and hypotheses testing using spacings. *Journal of Statistical Theory and Applications*, 8(2), 163-168.

[53] Willekens, F. (2002). Gompertz in context: the Gompertz and related distributions. *In E. Tabeau, A. van den Berg Jeths, & C. Heathcote (Eds.), Forecasting mortality in developed countries - insights from a statistical,demographic and epidemiological perspective, European studies of population*, 9,105-126. Dordrecht: Kluwer Academic Publishers.

[54] Wu, J.W, and Lee, W.C. (1999). Characterization of the mixtures of
      Gompertz distributions by conditional expectation of order statistics,
      Biometrical Journal, 41 , 371-381.

[55] Wu, S.J., Chang, C.T. and Tsai, T. R. (2003). Point and interval estima-
      tions for the Gompertz distribution under progressive type-II censor-
      ing. *Metron*, LXI(3), 403-418.

[56] Wu, J.W., Hung, W.L. and Tsai, C.H. (2004). Estimation of parameters
      of the Gompertz distribution using the least squares method. *Applied
      Mathematical Computation*, 158(1), 133-147.