

Research Article

An Iterated Local Search Algorithm for Estimating the Parameters of the Gamma/Gompertz Distribution

Behrouz Afshar-Nadjafi

Department of Industrial Engineering, Faculty of Industrial and Mechanical Engineering, Qazvin Branch, Islamic Azad University, Qazvin 34185-14161, Iran

Correspondence should be addressed to Behrouz Afshar-Nadjafi; afsharnb@alum.sharif.edu

Received 8 January 2014; Revised 7 May 2014; Accepted 12 May 2014; Published 5 June 2014

Academic Editor: Nikos D. Lagaros

Copyright © 2014 Behrouz Afshar-Nadjafi. This is an open access article distributed under the Creative Commons Attribution License, which permits unrestricted use, distribution, and reproduction in any medium, provided the original work is properly cited.

Extensive research has been devoted to the estimation of the parameters of frequently used distributions. However, little attention has been paid to estimation of parameters of Gamma/Gompertz distribution, which is often encountered in customer lifetime and mortality risks distribution literature. This distribution has three parameters. In this paper, we proposed an algorithm for estimating the parameters of Gamma/Gompertz distribution based on maximum likelihood estimation method. Iterated local search (ILS) is proposed to maximize likelihood function. Finally, the proposed approach is computationally tested using some numerical examples and results are analyzed.

1. Introduction

In probability and statistics, the Gompertz distribution is a continuous probability distribution. The Gompertz distribution was first introduced by Gompertz [1] to describe human mortality and establish actuarial tables. Since then, many investigators have used the Gompertz distribution or some related forms of it in a variety of studies. There are many forms of the Gompertz distribution in the literature. The Gompertz distribution is often applied to describe the distribution of adult lifespans by demographers [2, 3] and actuaries [4, 5]. Related fields of science such as biology [6] and gerontology [7] also considered the Gompertz distribution for the analysis of survival. More recently, computer scientists have also started to model the failure rates of computer codes by the Gompertz distribution [8]. In marketing science, it has been used as an individual-level model of customer lifetime [9]. Also, the Gompertz distribution is applied in reliability, life testing, epidemiological, and biomedical studies. Several such situations have been discussed by Ananda et al. [10], Walker and

Adham [11], Jaheen [12], and many others. The probability density function of the Gompertz distribution is

$$f(x) = b\eta e^{bx} e^{\eta} \exp\left(-\eta e^{bx}\right); \quad 0 \le x < \infty, \ b > 0, \ \eta > 0,$$
(1)

where *b* is the scale parameter and η is the shape parameter of the Gompertz distribution. In the actuarial and biological sciences and in demography, the Gompertz distribution is parameterized slightly and differently The Gompertz distribution is a flexible distribution that can be skewed to the right and to the left. The Gompertz density function can take on different shapes depending on the values of the shape parameter η . The *n*th moment of a Gompertz distributed random variable *X* is

$$E\left(X^{n}\right) = \frac{n!}{b^{n}}e^{n}E_{1}^{n-1}\left(\eta\right),$$
(2)

where $E_s^n(z) = (1/n!) \int_1^\infty (\ln x)^n x^{-s} e^{-zx} dz$ is the generalized integroexponential function [13]. When the shape



FIGURE 1: Three parameters Gamma/Gompertz distribution density function.

parameter η of a Gompertz distribution varies according to a Gamma distribution with shape parameter α and scale parameter β (mean = α/β), the distribution of *x* is Gamma/ Gompertz [9]. The Gamma/Gompertz distribution has been used as an aggregate-level model of customer lifetime and a model of mortality risks. Missov studied life expectancy resulting from a gamma-Gompertz force of mortality [14].

The probability density function of the Gamma/Gompertz distribution is

$$f(x) = \frac{bse^{bx}\beta^s}{(\beta - 1 + e^{bx})^{s+1}}; \quad 0 \le x < \infty, \ b > 0, \ \beta > 0, \ s > 0,$$
(3)

where *b* is the scale parameter and β and *s* are the shape parameters of the Gamma/Gompertz distribution (Figure 1). The cumulative distribution function of the Gamma/Gompertz distribution is

$$F(x) = 1 - \frac{\beta^{s}}{(\beta - 1 + e^{bx})^{s}}; \quad 0 \le x < \infty, \ b > 0, \ \beta > 0, \ s > 0.$$
(4)

When $\beta = 1$, this reduces to an exponential distribution with parameter *sb*:

$$F(x) = 1 - e^{-bsx}; \quad 0 \le x < \infty, \ b > 0, \ s > 0.$$
 (5)

The expected value of random variable *X* of a Gamma/Gompertz distribution depends on the values of *b*, β , and *s*; that is, we have $E(x) = (1/\beta)[\beta/(\beta - 1)] \ln \beta$ for *b*, s > 0, $\beta \neq 1$ and E(x) = 1/bs for *b*, s > 0, $\beta = 1$. Also, the median of this distribution is $(1/b) \ln\{\beta[(1/2)^{-1/s} - 1] + 1\}$.

Successful application of Gamma/Gompertz distribution depends on having acceptable statistical estimates of the distribution parameters. One of the best methods of obtaining a point estimator of a parameter is the method of maximum likelihood. As the name implies, the estimator will be the value of the parameter that maximizes the likelihood function. While the method of maximum likelihood is an excellent technique, sometimes complications arise in its use. For example, it may not always be possible to use calculus methods directly to determine the maximum of likelihood function. This difficulty is obvious in estimating the parameters of the Gamma/Gompertz distribution.

In this paper, we use iterated local search (ILS) to estimate Gamma/Gompertz parameters. The paper is organized as follows. Section 2 describes the notion of maximum likelihood estimation. Section 3 explains the basics of iterated local search (ILS). In Section 4, we explain the steps of ILS algorithm to solve the problem. Some numerical examples and their computational results are represented in Section 5. Finally, Section 6 contains the conclusions.

2. Maximum Likelihood Estimation

Maximum-likelihood estimation (MLE) is a method of estimating the parameters of a statistical model. In general, for a fixed set of data and underlying statistical model, the method of maximum likelihood selects values of the model parameters that produce a distribution that gives the observed data the greatest probability (i.e., parameters that maximize the likelihood function). In essence the method selects a set of model parameters that predicts that events that occur often in the data are very likely to occur and events that occur seldom in the data are predicted to occur with small probability. Maximum-likelihood estimation gives a unified approach to estimation, which is well-defined in the case of the normal distribution and many other problems.

Suppose that x is a random variable with probability distribution $f(x, \vec{\theta})$ where $\vec{\theta} = (\theta_1, \theta_2, \theta_3, \dots, \theta_k)$ is unknown parameters vector. Let x_1, x_2, \dots, x_n be the observed values in a random sample of size *n*. Then, the likelihood function of the sample is

$$L\left(\vec{\theta}\right) = f\left(x_{1};\vec{\theta}\right) = f\left(x_{2};\vec{\theta}\right)f\left(x_{3};\vec{\theta}\right)\cdots f\left(x_{n};\vec{\theta}\right)$$
$$= \prod_{i=1}^{n} f\left(x_{i};\vec{\theta}\right).$$
(6)

Note that the likelihood function is now a function of only the unknown parameters $\vec{\theta} = (\theta_1, \theta_2, \theta_3, \dots, \theta_k)$. The maximum likelihood estimators of $\vec{\theta}$ are the values that maximize the likelihood function $L(\vec{\theta})$. In practice, it is often more convenient to work with the logarithm of the likelihood function:

$$\operatorname{Ln} L\left(\vec{\theta}\right) = \operatorname{Ln} \prod_{i=1}^{n} f\left(x_{i}; \vec{\theta}\right).$$
(7)

The maximum likelihood estimators would be found by equating the k partial derivatives to zero and solving the resulting system of equations:

$$\frac{\partial}{\partial \vec{\theta}} L\left(\vec{\theta}\right) = \vec{0} \quad \text{or} \quad \frac{\partial}{\partial \vec{\theta}} \operatorname{Ln} L\left(\vec{\theta}\right) = \vec{0}. \tag{8}$$



ALGORITHM 1: Iterated local search (ILS) algorithm.

While the method of maximum likelihood is an excellent technique for many models and a maximum likelihood estimator can be found as an explicit function of the observed data x_1, \ldots, x_n , sometimes complications arise in its use. For example, it is not always easy to maximize the likelihood function because the equation(s) obtained from partial derivatives may be difficult to solve or no closed-form solution to the maximization problem is known or available. Therefore, an MLE has to be found numerically using optimization methods.

3. General Iterated Local Search

Iterated local search is a simple but powerful metaheuristic algorithm [15]. It applies local search to an initial solution until it finds a local optimum; then it perturbs the solution and restarts local search. The importance of the perturbation is obvious: a too small perturbation might not enable the system to escape from the basin of attraction of the local optimum just found. On the other side, a too strong perturbation would make the algorithm similar to a random restart local search.

A local search is effective if it is able to find good local optima, that is, if it can find the basin of attraction of those states. When the search space is wide and/or when the basin of attraction of good local optima is small, a simple multistart algorithm is almost useless. An effective search could be designed as a trajectory only in the set of local optima s^* , instead of in the set s of all the states.

The requirement on the perturbation of s is to produce a starting point for local search such that a local optimum different from s is reached. However, this new local optimum should be closer to s than a local optimum produced by a random restart. The *acceptance criterion* acts as a counter balance, as it filters and gives feedback to the perturbation action, depending on the characteristics of the new local optimum. A high level description of ILS steps is presented in Algorithm 1.

The design of ILS algorithms has several degrees of freedom in the choice of the initial solution, perturbation, and acceptance criteria.

The construction of initial solutions should be fast, and initial solutions should be a good starting point for local search. The fastest way of producing an initial solution is to generate it at random; however, this is the easiest way for problems that are unconstrained, whilst in other cases the construction of a feasible solution requires also constraint checking.

The perturbation is usually nondeterministic in order to avoid cycling. Its most important characteristic is the strength, roughly defined as the amount of changes made on the current solution. The strength parameter, λ , can be either fixed or variable. In the first case, the distance between perturbation and local search is kept constant, independently of the problem size. However, a variable strength is in general more effective, since it has been experimentally found that, in most of the problems, the bigger the problem size is, the larger should be the strength.

A second choice is the mechanism to perform perturbations. This may be a random mechanism, or the perturbation may be produced by a deterministic or semideterministic method.

The third important component is the acceptance criterion. Two extreme examples consist in (1) accepting the new local optimum only in case of improvement and (2) in always accepting the new solution. In-between, there are several possibilities. In this paper, we consider an annealing scheme for acceptance criterion. Simulated annealing is one of the most novel algorithms initially presented by Kirkpatrick et al. [16]. The algorithm is based upon that of Metropolis et al. which was originally proposed as a means of finding the equilibrium configuration of a collection of atoms at a given temperature [17]. Similar to other metaheuristic algorithms, it attempts to solve hard combinatorial optimization problems through controlled randomization. Sometimes simulated annealing was applied to parameters estimation [18].

The performance of ILS algorithm depends on the definition of the several control parameters. The choice of an appropriate λ is crucial for the performance of the algorithm. The value of parameter λ decrease during the search process; thus, at the beginning of the search, diversification is high and as it gradually goes on its search path, intensification becomes intense. Hereby, with the choice of appropriate λ , a dynamic balance is given between diversification and intensification.

4. Applying Hybrid ILS for Estimating Gamma/Gompertz Parameters

To estimate the three parameters of Gamma/Gompertz distribution, let x_1, \ldots, x_n be a random sample from the

Obtain a random sample from the distribution the size of which should be big enough. Initialize the ILS control parameter (λ , ε , C, τ_0)

select an initial solution $\theta_0 = (b, \beta)$;

Set $\theta = \theta_0$ $\overrightarrow{\theta^*} \leftarrow \text{Apply Local Search } (\vec{\theta})$ While termination condition not met do $\overrightarrow{\theta'} \leftarrow \text{Apply perturbation } (\overrightarrow{\theta^*})$

 $\vec{\theta}^{'*} \leftarrow \text{Apply Local Search } (\vec{\theta}^{'})$ $\vec{\theta}^{*} \leftarrow \text{Apply acceptance criterion } (\vec{\theta}^{*}, \vec{\theta}^{'*})$ Memorize Best Found Solution

ALGORITHM 2: Proposed hybrid iterated local search.

Gamma/Gompertz distribution. The log of the likelihood function is

End While

$$\ln L(b, \beta, s) = n \ln (b) + n \ln (s) + ns \ln (\beta) + b \sum_{i=1}^{n} x_i - (s+1) \ln \prod_{i=1}^{n} (\beta - 1 + e^{bx_i}).$$
⁽⁹⁾

When the derivatives are equated to zero, we obtain the following equations that must be solved to find the maximum likelihood estimators of *b*, β , and *s*:

$$\frac{\partial}{\partial b} \operatorname{Ln} L(b, \beta, s) = 0,$$

$$\frac{\partial}{\partial \beta} \operatorname{Ln} L(b, \beta, s) = 0,$$

$$\frac{\partial}{\partial s} \operatorname{Ln} L(b, \beta, s) = 0.$$
(10)
(11)

There is no closed form solution to these equations and it is very difficult to solve these equations using ordinary optimization techniques. To estimate the parameters *b*, β , and *s* we are to maximize $L(b, \beta, s)$ (or $\operatorname{Ln} L(b, \beta, s)$), using hybrid iterated local search.

For a given values of *b* and β , (11) gives an expression of Ln $L(b, \beta, s)$ as a function of only *s* which is called *reduced* Ln $L(b, \beta, s)$. Although Ln $L(b, \beta, s)$ is a nonconcave function, it is obvious from expression (11) that *reduced* Ln $L(b, \beta, s)$ is a concave function because second partial derivative of *reduced* Ln $L(b, \beta, s)$ with respect to *s* is nonpositive. Therefore, taking the first derivative of *reduced* Ln $L(b, \beta, s)$ with respect to *s* and setting $(\partial/\partial s)$ *reduced* Ln $L(b, \beta, s) = 0$ yield

$$s = \frac{n}{\operatorname{Ln}\left(\prod_{i=1}^{n} \left(\beta - 1 + e^{bx_i}\right)/\beta\right)}.$$
 (12)

Therefore, problem reduces to obtain parameters b and β ; s is computable from relation (12). Of course, only positive values for s are acceptable. To challenge this very difficult problem, we propose a hybrid iterated local search approach. In this regard, the steps of this algorithm are briefly presented in Algorithm 2, where the following notation is used:

$$\overrightarrow{\theta_0}$$
: Initial solution
 $\overrightarrow{\theta}$: Current solution
 $\overrightarrow{\theta'}$: Perturbed solution
 $\overrightarrow{\theta'}^*$: Local optimum solution
 $\overrightarrow{\theta'}^*$: Best solution
Ln $L(\overrightarrow{\theta})$: Value of the objective function at solution $\overrightarrow{\theta}$
 ε : Neighborhood step-length parameter
 λ : Perturbation step-length parameter
C: Geometric cooling factor

 τ_0 : Initial temperature.

The algorithm starts with an initial random solution for the problem (positive values for *b* and β), and by initializing the so-called perturbation step-length parameter λ , the neighborhood step-length parameter ε , cooling factor *C* and initial temperature τ_0 . At each solution, parameter *s* is calculated by relation (12) and only solutions are considered feasible for which *s* is positive. Two-neighborhood solution is generated by right shift and left shift of current solution by an amount of ε , for *b* or β , at random. Feasible improving direction is detected. The generated feasible solution in improving direction replaces the current one. Generating feasible neighborhood solution at same direction continues until a local

-300

-350

-400





-600-700

-800

FIGURE 2: Ln $L(b, \beta, s)$ versus iteration counter for Example 1: (a) n = 100, (b) n = 200, and (c) n = 300.

optimum solution ${\theta'}^{\hat{*}}$ is reached. This first local optimum solution sets as best solution θ^{*} . The cycle of algorithm repeats while the acceptance criterion is satisfied. At each repetition of the algorithm, a perturbed solution of the current local optimum solution θ'^* is generated as follows: with generating two binary random numbers z_1 and z_2 , we select a direction for perturbation (right or left) for each of the parameters b and β . Perturbed solution is obtained by adding to or subtracting from current local optimum solution, θ'^* , a dynamic amount, depending on the perturbation directions and λ , where λ is the perturbation step-length parameter playing an important role in our algorithm. The generated feasible solution replaces the current one. Local search procedure is applied to the newly chosen solution. We suppose that after the local optimum is reached, it is always acceptable (with checking positivity of *s*). After generating newly local optimum solution, θ'^* , a kind of annealing schedule is considered as acceptance criterion: accept all the improving new local optima and accept also the nonimproving ones with

a probability that is a function of the temperature τ and the difference of objective values, in formulas:

$$P(\text{Acceptance}) = \begin{cases} 1 & \text{if } \ln L\left(\overrightarrow{\theta'^*}\right) > \ln L\left(\overrightarrow{\theta'^*}\right) \\ \exp\left(-\frac{\ln L\left(\overrightarrow{\theta^*}\right) - \ln L\left(\overrightarrow{\theta'^*}\right)}{\tau}\right) \\ & \text{otherwise.} \end{cases}$$
(13)

It is obvious that control parameter, τ , is chosen with respect to the specific problem at hand. When adapting this general algorithm to a specific problem, the procedure to generate both initial and perturbed solutions is very important in addition to the control parameter. Hereby, the most important feature of this algorithm, as a metaheuristic, is the possibility of accepting a worst solution, which can allow it to prevent falling into local optimum trap. Obviously, the probability of



FIGURE 3: Ln $L(b, \beta, s)$ versus iteration counter for Example 2: (a) n = 100, (b) n = 200, and (c) n = 300.

accepting a worse solution decreases as the temperature decreases in each outer cycle. The initial temperature (τ_0) should be high enough that, in the first iteration of the algorithm, the probability of accepting a worse solution is, at least, of 80% [16]. The most commonly used temperature reducing function is geometric; that is, $\tau_i = C\tau_{i-1}$ in which C < 1 is constant. The termination condition happens when the system has reached to a desired energy level.

5. Numerical Examples

In this section, we solve two problem instances for illustrative and comparative purposes. In addition, to know the effects of the sample size, different sample sizes have been chosen and used to solve estimate parameters. We consider two Gamma/Gompertz distributions with parameters b = 2, $\beta =$ 4, and s = 3 (Example 1) and b = 3, $\beta = 3$, and s = 3(Example 2), respectively. Let x_1, x_2, \ldots, x_n be the observed values from each distribution in a random sample of size n = 100, 200, and 300. The coding has been done using MatLab 7.12 and run on an Intel Core i5 CPU processor with 4 GB of RAM.

The performance of metaheuristic algorithms depends excessively on the value of their parameters. To know the effects of the perturbation rule and acceptance criterion, different approaches have been chosen and used to estimate parameters. We obtained best results with perturbation rule according to $\lambda = (101-i) * \varepsilon$ (in which *i* is iteration counter) and acceptance criterion based on $\tau > \tau_0$.

In this paper, the proposed algorithm parameter values are selected through the computational experiments. The experiments indicated that the best values for the parameters of the algorithms in our examples are as follows: C = 0.975 and $\tau_0 = 100$. $\varepsilon = 0.01$ is selected for the examples.

Tables 1 and 2 show the results. Figures 2 and 3 show the likelihood function for all the values evaluated during the process of maximization. The experimental results demonstrate that calibration provides high quality solutions. It is quite clear that as the sample size increases the estimation will

Example 1	$(b, \beta, s) = (2, 4, 3)$		
Ν	100	200	300
$(\hat{b}, \hat{\beta}, \hat{s})$	(2.02, 4.12, 3.0728)	(1.98, 4.06, 2.944)	(1.98, 3.93, 2.9506)
$\operatorname{Ln} L(\widehat{b}, \widehat{\beta}, \widehat{s})$	-318.6441	-645.8559	-948.7620
Run time (sec)	2.19	12.38	19.06

TABLE 1: Experimental result for Example 1.

TABLE 2: Experimental result for Example 2.

Example 1	$(b, \beta, s) = (3, 3, 3)$		
n	100	200	300
$(\widehat{b},\widehat{eta},\widehat{s})$	(2.87, 3.06, 2.8466)	(2.85, 2.91, 3.0412)	(3.09, 2.93, 3.029)
$\operatorname{Ln} L(\widehat{b}, \widehat{\beta}, \widehat{s})$	-233.1599	-458.5367	-651.2425
Run time (sec)	1.84	9.76	23.05

be better. It is straightforward from the primary estimation theory that the bigger the sample size the better the estimation. With comparing the CPU time requirements, it is clear that as the sample size increases more run time is needed. It is because as the sample size increases the more complicated will be the likelihood function to maximize.

6. Conclusions

We considered the Gamma/Gompertz distribution, which is often encountered in customer lifetime and mortality risks distribution literature. We developed hybrid iterated local search to estimation of Gamma/Gompertz parameters. To improve the efficiency of the proposed algorithm, its parameters were selected through the computational experiments. The results of the computational experiment showed that calibration provides high quality solutions. The large size samples perform better than the small size ones, measured by the solution accuracy index but with more CPU time.

Conflict of Interests

The author declares that there is no conflict of interests regarding the publication of this paper.

References

- B. Gompertz, "On the nature of the function expressive of the law of human mortality and on the new mode of determining the value of life contingencies," *Philosophical Transactions of the Royal Society A*, vol. 115, pp. 513–580, 1825.
- [2] J. W. Vaupel, "How change in age-specific mortality affects life expectancy," *Population Studies*, vol. 40, no. 1, pp. 147–157, 1986.
- [3] S. H. Preston, P. Heuveline, and M. Guillot, *Demography: Measuring and Modeling Population Processes*, Blackwell, Oxford, UK, 2001.
- [4] B. Benjamin, H. W. Haycocks, and J. H. Pollard, *The Analysis of Mortality and Other Actuarial Statistics*, Heinemann, London, UK, 1980.
- [5] W. J. Willemse and H. Koppelaar, "Knowledge elicitation of Gompertz' law of mortality," *Scandinavian Actuarial Journal*, vol. 2, pp. 168–179, 2000.

- [6] A. C. Economos, "Rate of aging, rate of dying and the mechanism of mortality," *Archives of Gerontology and Geriatrics*, vol. 1, no. 1, pp. 3–27, 1982.
- [7] K. S. Brown and W. F. Forbes, "A mathematical model of aging processes," *Journals of Gerontology*, vol. 29, no. 1, pp. 46–51, 1974.
- [8] K. Ohishi, H. Okamura, and T. Dohi, "Gompertz software reliability model: estimation algorithm and empirical validation," *Journal of Systems and Software*, vol. 82, no. 3, pp. 535–543, 2009.
- [9] A. C. Bemmaor and N. Glady, "Modeling purchasing behavior with sudden "death": a flexible customer lifetime model," *Management Science*, vol. 58, no. 5, pp. 1012–1021, 2012.
- [10] M. M. Ananda, R. J. Dalpatadu, and A. K. Singh, "Adaptive bayes estimators for parameters of the gompertz survival model," *Applied Mathematics and Computation*, vol. 75, no. 2-3, pp. 167–177, 1996.
- [11] S. G. Walker and S. A. Adham, "A multivariate Gompertz-type distribution," *Journal of Applied Statistics*, vol. 28, pp. 1051–1065, 2001.
- [12] Z. F. Jaheen, "A Bayesian analysis of record statistics from the Gompertz model," *Applied Mathematics and Computation*, vol. 145, no. 2-3, pp. 307–320, 2003.
- [13] M. Milgram, "The generalized integro-exponential function," *Mathematics of Computation*, vol. 44, no. 170, pp. 443–458, 1985.
- [14] T. I. Missov, "Gamma-Gompertz life expectancy at birth," Demographic Research, vol. 28, pp. 259–270, 2013.
- [15] H. R. Lourenco, O. Martin, and T. Stutzle, "A beginner's introduction to iterated local search," in *Proceeding of Meta-heuristics International Conference (MIC '01)*, vol. 1, pp. 1–6, Porto, Portugal, 2001.
- [16] S. Kirkpatrick, C. D. Gelatt Jr., and M. P. Vecchi, "Optimization by simulated annealing," *Science*, vol. 220, no. 4598, pp. 671–680, 1983.
- [17] N. Metropolis, A. W. Rosenbluth, M. N. Rosenbluth, A. H. Teller, and E. Teller, "Equation of state calculations by fast computing machines," *The Journal of Chemical Physics*, vol. 21, no. 6, pp. 1087–1092, 1953.
- [18] B. Abbasi, A. H. Eshragh Jahromi, J. Arkat, and M. Hosseinkouchack, "Estimating the parameters of Weibull distribution using simulated annealing algorithm," *Applied Mathematics and Computation*, vol. 183, no. 1, pp. 85–93, 2006.

