JOURNAL OF APPLIED QUANTITATIVE METHODS

Quantitative Methods Inquires

INFORMATION THEORETIC ESTIMATION IMPROVEMENT TO THE NONLINEAR GOMPERTZ'S MODEL BASED ON RANKED SET SAMPLING

Enrico CIAVOLINO
Università Salento, Salento, Italy

E-mail: enrico.ciavolino@ateneo.unile.it

Amajd D. AL-NASSER
Yarmouk University, Irbid, Jordan

E-mail: amjadyu@yahoo.com

Abstract: *The aim of this paper is to apply both Generalized Maximum Entropy (GME) estimation method and Ranked Set Sampling (RSS) technique to improve the estimations of the Gompertz's Model. The Gompertz's model is a simple formula which expresses the geometrical relationship between the force of mortality and age.*
The choice of evaluating the RSS is due to the fact that in many practical applications of the Gompertz's model, as in biological or environmental sciences, the variable of interest is more costly to measure but is associated with several other easily obtainable.
In this paper, we have used Monte Carlo experiments to illustrate the performance of the GME estimator based on two different sampling techniques: the Simple Random Sample (SRS) and RSS. Moreover, the results are compared with the traditional Maximum Likelihood Estimates (MLE).

Key words: *Generalized Maximum Entropy; Ranked Set Sampling; Gompertz's Model; Maximum Likelihood Estimates; Monte Carlo experiments*

1. Introduction

Gompertz's law (Gompertz, 1865) has a very important role in modelling the human mortality rate and providing the actuarial tables. Moreover, in recent years, it has been applied to several other fields, such as in the fertility rate model (Booth, 1984), in biological medical data (Ricklefs et al., 2002), in environmental data, or in the study of reliability (Yamada et al., 1985).

In this paper, we propose the GME estimator as an alternative to Maximum Likelihood Estimation (MLE), since it is widely used in literature in cases of non linear function (Golan et al., 2001). In particular we can highlight some points:

- The GME approach uses all the data points and does not require restrictive moment or distributional error assumptions.
- Thus, unlike the MLE estimator, the GME is robust for a general class of error distributions.
- The GME estimator can be used when the sample is small, when there are many covariates, and when the covariates are highly correlated.

It means that the GME was proposed because it is particularly useful when there are no restrictive sampling assumptions or in the case of ill-posed problems or

underdetermined ones, and also because in these cases MLE is unattractive and not robust to the underlying (unknown) distribution (Golan, 2008).

Moreover, RSS (McIntyre, 1952) could be used to increase the precision of estimates and to reduce costs, by using the researchers' experiences or inexpensive measurements.

The format of this paper is as follows. In Section 2 the Gompertz's model is introduced in its characteristics and analytical formulation by considering the MLE method. In Section 3, firstly the classic GME formulation is specified and then the Gompertz's model is expressed in the framework of the GME. In Section 4, the RSS method is presented and a discussion on how it can be utilized for applications in which the Gompertz's model is used. In Section 5, Monte Carlo experiments show the numerical performance for the proposed estimators, based on the RSS and SRS sampling schemes. Section 6 contains concluding remarks and discussion of future works.

2. The Gompertz's Model

Benjamin Gompertz, in 1825, showed that the mortality rate increases in a geometrical progression, defining one of the most informative actuarial functions for investigating the ageing process. He observed a law of geometrical progression in death rates by analysing a sample of people, aged between 20 and 60 years, in England, Sweden, and France. Gompertz's law has become the most successful law to model the dying out process of living organisms (Willemse et al., 2000). The relationships of the force of mortality, such as the power function of age, is also called the Weibull model (Carey, 2001) and it can be expressed by the following formula:

$$y = B \cdot C^x + \varepsilon \tag{1}$$

Where $B>0$ reflects the general level of mortality in the population, and $C>1$ is Gompertz's constant that reflects the rate at which the force of mortality increases with age; $x>0$ is the *age*, and y is the *force of mortality* at *age* x.

The estimates of the parameters B and C can be obtained numerically using different methods, however, the MLE is the most commonly used for this model and it is discussed by Garg *et al.* (1970).

The MLE estimation method, presents the assumption that the random variables D_x *(number of deaths at age x before reaching age x+1)* follows a Binomial distribution: $B(n_x, q_x)$. The p_x, reported in the following formula (2), is the conditional probability function of the Gompertz's model given in (1):

$$p_x = e^{-\int_x^{x+1} \mu_t dt} = e^{-\int_x^{x+1} BC^t dt} = e^{-BC^x (C-1)/\log C} \tag{2}$$

Or considering the log function:

$$\log p_x = -B \cdot C^x \cdot (C-1)/\log C \text{ and } \log p_{x+1} / \log p_x = C \tag{3}$$

The other parameters are defined as follows: $q_x=1-p_x$, n_x is the number of people of age x and d_x is the value of the random variable D_x. Given this assumption, the logarithm of the joint likelihood function is the following:

$$\log L(B,C) = \sum_x (n_x - d_x) \cdot \log p_x + d_x \cdot \log(1 - p_x) \tag{4}$$

The general way of estimating the parameters is to use Nonlinear Maximization (NM), so as to find the minimum of the following function and to plug into (2) for the p_x:

$$\min_{B,C} \sum_x \left[\left(n_x - d_x\right) \cdot B \cdot C^x \cdot (C-1)/\log C - d_x \cdot \log\left(1 - e^{-B \cdot C^x \cdot (C-1)/\log C}\right) \right] \qquad (5)$$

The NM problem can be solved via numerical methods such as the Newton-Raphson iteration or the Simplex algorithm.

3. Generalized Maximum Entropy

In the framework of the Information Theoretic we propose the use of the Generalized Maximum Entropy (GME) to estimate the non linear relationship between age and the force of mortality. In this section, we start by briefly describing the traditional maximum entropy (ME) estimation method, then we introduce the GME formulation as a method for recovering information from the Gompertz's model.

3.1. The GME estimation method

The *Entropy of Information* was first introduced by Claude Shannon in 1948 as a propriety associated to any probability distribution, defining an axiomatic method of measuring the uncertainty (state of knowledge) of a collection of events.

Letting *X* be a random variable with possible outcomes $\{x_1, x_2,\ldots, x_s\}$, with the corresponding probability $P = \{p_1, p_2,\ldots, p_s\}$ such that $\Sigma p_i=1$, Shannon defined the *entropy of information* of probabilities distribution function as:

$$H(P) = -k \cdot \sum_{i=1}^{s} p_i \cdot \log(p_i) \qquad (6)$$

Where *k* is a constant usually equal to 1; and $0 \cdot \ln(0)=0$. The quantity $\{- log(p_i)\}$ is called self information of the x_i event. The average on the self information is defined as the Entropy. The function *H(P)* is called *Entropy*, *Shannon's Entropy* or *Information Entropy*.

Edwin Thompson Jaynes (1957a, 1957b) extended the entropy of information by defining the *Maximum Entropy Principle* (MEP). The MEP estimates an unknown probability distribution from given moment constraints and adding up normalization constraints on the probabilities.

The frequency that maximizes entropy is an intuitively reasonable estimate of the true distribution when we lack any other information. If we have information about an experiment, such as the sample moments, or non-sample information about the random variable, such as restrictions from economic theory, we can to alter our "intuitively reasonable" estimate. The method of maximum entropy proceeds by choosing the distribution that maximizes entropy, subject to the sample and non-sample information.

Letting *X* be a random variable with possible outcomes $\{x_1, x_2,\ldots, x_s\}$, the objective of the MEP is to recover the unknown probability distribution $P = \{p_1, p_2,\ldots, p_s\}$, taking into account: *consistency constraints*, by defining the functions $\{f_1(x_i), f_2(x_i), \ldots, f_T(x_i)\}$, which represents the constraints (or information) generated by the data; *measurable values* $\{y_1, y_2, \ldots, y_T\}$, obtained by a priori knowledge on the phenomenon, or by training dataset.

The constraints generated from the data and the measurable values, are expressed by the following equation:

$$\sum_{i=1}^{s} f_t(x_i) \cdot p_i = y_t \quad t = 1,2,\ldots,T \qquad (7)$$

The adding up *normalization constraints*, which means $\Sigma p_i=1$, are the following:

$$\sum_{i=1}^{s} p_i = 1 \qquad (8)$$

The MEP estimates the probability distribution by the maximization of the entropy function, equation (6), based on the consistency and normalization constraints, equation (7, 8), where the consistency constraints are expressed in form of expectation values of a discrete random variable[1]:

Amos Golan *et al.* proposed in 1996 an alternative method to solve many standard and ill-posed econometric problems in the framework of the MEP developed by Jaynes,

This information theoretic method, which is called Generalized Maximum Entropy (GME), is based on the *re-parameterization* and *re-formulation* of a general linear model. Considering a regression model $\mathbf{y}=\mathbf{X}\boldsymbol{\beta}+\boldsymbol{\varepsilon}$ with *n* units and *m* variables, the coefficients and the error terms can be re-parameterized as a convex combination of expected value of a discrete random variable, as in the

following equation:

$$\mathbf{y}_{n,1} = \mathbf{X}_{n,m} \cdot \boldsymbol{\beta}_{m,1} + \boldsymbol{\varepsilon}_{n,1} = \mathbf{X}_{n,m} \cdot \mathbf{Z}_{m,m \cdot M} \cdot \mathbf{p}_{m \cdot M,1} + \mathbf{V}_{n,n \cdot N} \cdot \mathbf{w}_{n \cdot N,1} \qquad (9)$$

The matrix inner products ($\mathbf{Z}\cdot\mathbf{p}$) and ($\mathbf{V}\cdot\mathbf{w}$), represent respectively the re-parameterization of the regression coefficients and the error terms as in form of expected value of a discrete random variable.

The matrices **Z** and **V** are diagonal and the generic matrix element is represented respectively by the vectors $\mathbf{z}_k' = [-c \quad -c/2 \quad 0 \quad c/2 \quad c]$ with {k = 1, . . . , m} and $\mathbf{v}_h' = [-b \quad -b/2 \quad 0 \quad b/2 \quad b]$ with {h = 1, . . . , n}.

These vectors ($\mathbf{z}_k$ and $\mathbf{v}_h$) define the support values, called *fixed points*, usually with five elements (M=N=5) with a given constants '*c*' and '*b*', uniformly and symmetrically chosen around zero with equally spaced distance discrete points, or as Golan suggested, for the error terms, to follow the six-sigma rule (Pukelshiem, 1994), where the number of fixed points is 3, and the constant c is equal to $\hat{s}_y$.

The super vectors **p** and **w** associated are probabilities and have to be estimated by maximization of the Shannon entropy function:

$$H(p,w) = -\mathbf{p}_{1,m \cdot M}' \cdot \ln \mathbf{p}_{m \cdot M,1} - \mathbf{w}_{1,n \cdot N}' \cdot \ln \mathbf{w}_{n \cdot N,1}$$

subjected to some *normalization* and *consistency constraints*. The steps for the GME algorithm are shown in the following table 1.

Table 1. The Generalized Maximum Entropy Algorithm

1. **Re-parameterize** the unknown parameters and the disturbance terms as a convex combination *of expected value of a discrete random variable;*
2. **Re-formulate the model** with the new re-parameterization as the data constraint;
3. **Define the GME problem as non-linear programming** problem in the following form:

Objective Function = Shannon's Entropy Function

1. The consistency constraints, which represents the new formulation of the model;
2. The normalization Constraints.

4. **Solve the non-linear programming** by using numerical method

The constraints defined for estimating the unknown parameters refer to *consistency and normalization constraints*. The first one represents the information generated from the

data, that means a part of the model defined in the equation (9); the second one identifies the conditions: $0 \le p_{kj} \le 1$, {j=1, …,M; k=1,…,m}, $\sum p_{kj}=1$ {k=1,…,m} and $0 \le w_{hj} \le 1$, {j=1,…,N; h=1,…,n}, $\sum w_{hj}=1$ {h=1,…,n}.

The main advantages of using GME estimation method, as above defined (*supra* § 1), are its desirable properties which can be briefly summarized: does not require restrictive moments or distributional error assumptions; it's robust for a general class of error distributions; may be used with small samples, with many highly correlated covariates; moreover, using the GME method, it is easy to impose nonlinear and inequality constraints.

Therefore the GME works well in case of ill-behaved data and the above listed cases, where the MLE estimator cannot proceed.

In the following section the GME formulation for the Gompertz's model is explained, moreover the definition of the optimization function and both consistency and normalization constraints will be discussed.

3.2. GME for the Gompertz's Model

The GME formulation as a method for recovering information from the Gompertz's model, starts from the relationship between the age and the force of mortality, expressed by:

$$\mathbf{y}_{n,1} = B \cdot C^{\mathbf{x}_{n,1}} + \boldsymbol{\varepsilon}_{n,1} \tag{9}$$

The starting point, following the algorithm in table 1, is the *re-parameterization* of the unknown parameters and error terms, as a convex combination of excepted values. For the Gompertz's model, the number of predictor variables is just one, age, which means, as in the above general formulation, m=1 and the matrix **Z** is just a vector $\mathbf{z}_k' = [-c \quad -c/2 \quad 0 \quad c/2 \quad c]$, considering M=5 fixed points. The error term is obtained by considering the diagonal matrix with generic element $\mathbf{v}_k' = [-3 \cdot \hat{s}_y \quad 0 \quad 3 \cdot \hat{s}_y]$, based on the three-sigma rule, with N=3 fixed points.

The *re-parameterization* as convex combination of expected values is expressed in the following equations:

$$B = \mathbf{z}_{1,M}' \cdot \mathbf{p}_{M,1} \tag{10}$$

$$C = \mathbf{c}_{1,M}' \cdot \mathbf{q}_{M,1} \tag{11}$$

$$\boldsymbol{\varepsilon}_{n,1} = \mathbf{V}_{n,n \cdot N} \cdot \mathbf{w}_{n \cdot N,1} \tag{12}$$

Moreover, *normalization constraints* are necessary, because for each probability vector of the coefficients and error terms (ex., $\mathbf{p}_{M,1}$), the sum of probabilities estimated have to be equal to 1, which means $\sum p_j=1$, $\sum q_j=1$ and $\sum w_{hj}=1$ {h=1, …, n}.

These constraints are formalized by the following expressions:

$$\mathbf{p}_{1,M}' \cdot \mathbf{1}_{M,1} = 1 \tag{13}$$

$$\mathbf{q}_{1,M}' \cdot \mathbf{1}_{M,1} = 1 \tag{14}$$

$$\mathbf{J}_{n,n \cdot N}^{*} \cdot \mathbf{w}_{n \cdot N,1} = \mathbf{1}_{n,1} \tag{15}$$

The matrix **J*** is the Kronecker product between the identity matrices $\mathbf{I}_{n,n}$ and the vector of one $\mathbf{1}_{M,1}$, which means $\mathbf{J}_{n,n \cdot N}^{*} = (\mathbf{I}_{n,n} \otimes \mathbf{1}_{1,N}')$:

$$\mathbf{J}^{*}_{n,n\cdot N}\cdot\mathbf{w}_{n\cdot N,1}=\mathbf{1}_{n,1}\quad=>\begin{bmatrix}111 & \mathbf{0} & \cdots & \mathbf{0}\\ \mathbf{0} & 111 & \cdots & \mathbf{0}\\ \vdots & \vdots & \ddots & \vdots\\ \mathbf{0} & \mathbf{0} & \cdots & 111_{n}\end{bmatrix}\cdot\begin{bmatrix}\mathbf{w}_{1}\\ \mathbf{w}_{2}\\ \vdots\\ \mathbf{w}_{n}\end{bmatrix}=\begin{bmatrix}1\\ 1\\ \vdots\\ 1_{n}\end{bmatrix} \tag{16}$$

Given the re-parameterization, the Gompertz's model can be *re-formulated* in the following way, where all the parameters, unknowns and error terms, are expressed as expected values:

$$\mathbf{y}_{n,1}=\left(\mathbf{z}^{'}_{1,M}\cdot\mathbf{p}_{M,1}\right)\cdot\left(\mathbf{c}^{'}_{1,M}\cdot\mathbf{q}_{M,1}\right)^{\mathbf{x}_{n,1}}+\mathbf{V}_{n,n\cdot N}\cdot\mathbf{w}_{n\cdot N,1} \tag{17}$$

The definition of *GME problem as non linear programming problem* is the following:

Objective function: Max $H(P,Q,W)=-\mathbf{p}^{'}\cdot\ln\mathbf{p}-\mathbf{q}^{'}\cdot\ln\mathbf{q}-\mathbf{w}^{'}\cdot\ln\mathbf{w}$

Consistency Constraints: $\mathbf{y}_{n,1}=\left(\mathbf{z}^{'}_{1,M}\cdot\mathbf{p}_{M,1}\right)\cdot\left(\mathbf{c}^{'}_{1,M}\cdot\mathbf{q}_{M,1}\right)^{\mathbf{x}_{n,1}}+\mathbf{V}_{n,n\cdot S}\cdot\mathbf{w}_{n\cdot S,1}$

Normalization Constraints: $\mathbf{p}^{'}_{1,M}\cdot\mathbf{1}_{M,1}=1\ ;\ \mathbf{q}^{'}_{1,M}\cdot\mathbf{1}_{M,1}=1\ ;\ \mathbf{J}^{*}_{n,n\cdot N}\cdot\mathbf{w}_{n\cdot N,1}=\mathbf{1}_{n,1}$

The non-linear programming system is solved by the formulation of the Lagrangian function and the first order conditions which provides the basis for the solution. The Lagrangian function is expressed by the following formulation:

$$L=-\mathbf{p}^{'}\cdot\ln\mathbf{p}-\mathbf{q}^{'}\cdot\ln\mathbf{q}-\mathbf{w}^{'}\cdot\ln\mathbf{w}+\boldsymbol{\lambda}^{'}\cdot\left[\mathbf{y}-\left(\mathbf{z}^{'}\cdot\mathbf{p}\right)\cdot\left(\mathbf{c}^{'}\cdot\mathbf{q}\right)^{\mathbf{x}}-\mathbf{V}\cdot\mathbf{w}\right]+$$
$$+\theta\cdot\left[1-\mathbf{p}^{'}\cdot\mathbf{1}\right]+\nu\cdot\left[1-\mathbf{q}^{'}\cdot\mathbf{1}\right]+\boldsymbol{\tau}^{'}\cdot\left[\mathbf{1}-\mathbf{J}^{*}\cdot\mathbf{w}\right] \tag{18}$$

where θ, ν, $\boldsymbol{\lambda}$, $\boldsymbol{\tau}$, are respectively the scalars and the vectors of the Lagrangian multipliers. By taking the gradient of L it is possible to derive the first-order-condition. However, the equations system will not be in closed form and to get the final values, a numerical optimization technique (successive quadratic programming method) should be used to compute probabilities (Ciavolino, 2007).

The estimations can be expressed by:

$$\hat{B}=\mathbf{z}^{'}_{1,M}\cdot\hat{\mathbf{p}}_{M,1} \tag{19}$$

$$\hat{C}=\mathbf{c}^{'}_{1,M}\cdot\hat{\mathbf{q}}_{M,1} \tag{20}$$

$$\hat{\varepsilon}=\mathbf{V}_{n,n\cdot S}\cdot\hat{\mathbf{w}}_{n\cdot S,1} \tag{21}$$

Respectively for the general level of mortality (19), the force of mortality (20) and the error term (21).

4. Ranked Set Sampling Method

The concept of Ranked Set Sampling is a recent development that enables more structure to be provided to the collected sample items, although the name is a bit of a misnomer as it is not as much a sampling technique as a data measurement technique.

This approach to data collection was first proposed by McIntyre in 1952, for situations where taking the actual measurements for sample observations is difficult (e.g., costly, destructive, time-consuming), but mechanisms for either informally or formally ranking a set of sample units is relatively easy and reliable.

In particular, McIntyre was interested in improving the precision in estimation of average yield from large plots of arable crops without a substantial increase in the number of fields from which detailed, expensive and tedious measurements needed to be collected. For discussions of some of the settings where ranked set sampling techniques have found application, see Chen *et al.* (2004).

The idea of using the RSS for the Gompertz's model, is in some practical applications, where the response variable is too expensive to measure , destructive or time-consuming, but but the predictor variable can be measured easily with relatively negligible cost.

Some examples can be found in studies in the ecological or biology fields, where, for instance, the death rate of plants or insects caused by the contamination in hazardous waste sites are analyzed using the Gompertz's law. The analysis can be improved by inspections for ranking areas of soil based on visual inspection or other expert opinion about the sample units (Chen *et al.*, 2004).

The scheme we used in this study is the balanced RSS, which involves drawing *m* sets of Simple Random Samples, each of size *m* from a population, and ranking each set with respect to the variable of interest, for instance *the age*. Then, from the first set, the element with the smallest rank is chosen for the actual measurement. From the second set, the element with the second smallest rank is chosen. The process is continued until we have selected the largest unit from the last simple random sample.

The procedure can be formalized as the following steps:

1. Select m^2 units randomly from the population.
2. Randomly allocate the m^2 units into *m* subsets, each of size m.
3. Order the units within each subset, based on the perception of the interest variable.
4. In the RSS, the smallest unit in the first subset is selected for actual measurement, the second smallest unit in the second subset is selected for actual measurement, we continue in this process until the largest ranked unit is selected from the m^{th} subset.
5. These four steps are called *cycle* and can be repeated *r* times until the desired sample size is reached, equal to $n = r \cdot m$.

If just three ranks and one cycle are considered, the selected RSS is denoted by: $\{X_{[1]}, X_{[2]}, X_{[3]}\}$. In order to select a RSS with a sample size equal to $n = r \cdot m$, the cycle is repeated r independent times, yielding the following sample:

$$\left\{X_{[1],j}, X_{[2],j}, ..., X_{[m],j}\right\}, for \quad j = 1, ..., r \tag{22}$$

It can be noted that the selected elements are mutually independent order statistics but not identically distributed. In practice, the set size *m* is kept small to ease the visual ranking, RSS literature suggested that m = 2, 3, 4, 5 or 6.

Table 2. Ranked Set Sampling Design

	Rank (m=3)		
Cycle (r=4)	**Young**	**Adult**	**Elderly**
1	$X_{[1],1}$	-	-
	-	$X_{[2],1}$	-
	-	-	$X_{[3],1}$
2	$X_{[1],2}$	-	-
	-	$X_{[2],2}$	-
	-	-	$X_{[3],2}$
3	$X_{[1],3}$	-	-
	-	$X_{[2],3}$	-

	-	-	$X_{[3],3}$
4	$X_{[1],4}$	-	-
	-	$X_{[2],4}$	-
	-	-	$X_{[3],4}$

Table 2 shows an illustrative example, where the set size m is equal to 3 (the number of ranks) and the cycles r = 4. In the table, we assume that the units are selected according to their age, therefore there are three ranks: *Young*, *Adult*, *Elderly*, and the rows represent the ordered sample within each cycle, where $X_{[i],i}$ is the sample unit included in the RSS, which represents the i^{th} order unit in the i^{th} cycle.

The number of units randomly selected are 36, $m^2=3^2=9$, for 4 cycles, but only 12 are included in the RSS.

5. Simulation and Results

We perform different simulation studies in order to draw conclusions about the performance of GME and MLE estimation methods for the Gompertz's model by using the RSS technique. The simulation experiments start with the analysis of the fixed points to evaluate the sensitivity of the GME for the Gompertz's model. The analysis of the sensitivity is made by changing the value of the constant 'c' and the number of the fixed points. This analysis is recommended and useful (Golan *et al.*, 1996), to verify the support spaces on the coefficients and error terms to measure the sensitivity of results across support space specifications.

Four simulation experiments are considered to measure the sensitivity of the GME estimator: the first two simulations, to determine the variance and the bias of the Gompertz's coefficients; the last two simulations, to examine the variance and the bias of the error terms.

The sensitivity analysis is measured in term of Bias and Mean Squared Error (MSE), as in the following equations for the B coefficient:

$$Bias(B) = n^{-1} \cdot \sum_{i=1}^{n} \left(\hat{B}_i \right) - B$$

$$MSE(B) = n^{-1} \cdot \sum_{i=1}^{n} \left(\hat{B}_i - B_i \right)^2$$

After the sensitivity simulation studies, that allow to choose the best support space specification for the GME estimator, we performed two simulation studies for the evaluation GME and MLE performance based on the SRS and RSS.

The measures Bias and MSE are used to evaluate the performance of the estimator applied to both sampling techniques, computing also relative efficiency (*eff*); where the *eff* of B is defined as:

$$eff(B) = \frac{MSE(B_{GME})}{MSE(B_{MLE})}$$

To perform all the simulation studies, the parameters of the Gompertz's function B and C were first initialized to 0,3 and 2. Using these values as initial values, a simulation study was carried out by generating 1000 samples according to the following relationship:

$$y_i = 0,3 \cdot (2)^{x_i} + \varepsilon_i \quad , i = 1,2,...,n \qquad (23)$$

where $\varepsilon \sim N(0,1)$ and $X \sim Exp(1)$. The sensitivity simulation studies for the GME parameters choice were conducted by fixing the sample size to be 20.

All the pseudo distributions varieties are generated from a build in subroutines of International Mathematical and Statistical Library (IMSL). The solution of GME system were solved and generated by using a successive quadratic programming method to solve a nonlinear programming problems based on the NCONF and NLPQL subroutines. The basic FORTRAN codes for computational purposes were developed by Schittkowski (1986) and necessary modifications were made by the authors.

5.1. The Sensitivity Analysis of the GME Estimates

Tables 3, 4, 5 and 6 show the results of the sensitivity analysis for the *fixed points*. It has been considered four experiments: the first two to evaluate the support value 'c' and the number of fixed points for the parameters B and C; the last two for the error term (*supra* § 3).

Experiment 1. This experiment is performed to select the support parameter bounds of B and C. we start by using only three support values for each of these parameters in the interval [-c, 0, c], where c = 1, 5, 10, 50 or 100. The support bounds of the error term are fixed to be three data points selected according to the three sigma rule. The results of this simulation study are given in Table.3. The results indicated that the GME estimates were more accurate and more efficient than the MLE estimates for all the selected support bound. However, the best support bound should be selected in the interval [-5, 0, 5] for both parameters, which; almost, gives the minimum MSE and bias for the GME estimates.

Table 3. Selecting the 'c' value of the parameters (B e C) support points

	METHODS	GME: different 'c' VALUES of the fixed points					MLE
PARAMETERS		[-1, 1]	[-5, 5]	[-10, 10]	[-50, 50]	[-100, 100]	
B	***Bias***	-0,0098	-0,0107	-0,0097	-0,0114	-0,012	0,1715
	MSE	0,0108	0,011	0,0115	0,0153	0,0131	0,0394
C	***Bias***	-0,0159	-0,0029	-0,0092	0,0171	0,0141	0,0504
	MSE	0,0028	0,0022	0,0023	0,0026	0,0028	0,0034

Experiment 2. This experiment is repeated under the same assumptions of experiment two, and by fixing the support bounds of the parameters in the interval [-5, 5]. The aim now is to determine the number of support points within the suggested interval that leads to better results. Consequently, we start to increase the numbers of support points and allocate them in an equidistant fashion. The results of this experiment in Table. 4 suggested that there is no statistical improvement of the estimators if we increase the number of support points from 3 to 7 data points. Therefore, according to this simulation experiment, we suggest to fix the number of support points to be three in the interval [-5, 5].

Table 4. Fixing the number of support points for the parameters (B e C)

	METHODS	GME: Different NUMBERS of fixed points in [-5 5]					MLE
PARAMETERS		3	4	5	6	7	
B	***Bias***	-0,0107	-0,0071	-0,0104	-0,0077	-0,0098	0,1715
	MSE	0,011	0,0128	0,0118	0,014	0,0128	0,0394
C	***Bias***	-0,0029	-0,0136	-0,0112	-0,008	-0,0077	0,0504
	MSE	0,0022	0,0024	0,0023	0,0024	0,0021	0,0034

Experiment 3. We conducted a sampling experiment based on the experimental design outlined above and based on the results of experiment 1 and experiment 2. Hereafter, the support bounds of the error term is selected in the interval [-c S_y, 0, c S_y], then we start changing the value of 'c' to be equal to 1, 2, 3, 4 and 5. The results of this experiment in Table 5 indicate that the best value for 'c' is 3, this result is consistent with the results of Golan et al (1996), which suggest in using the three sigma rules when we setup the support bonds of the error terms.

Table 5. Selecting the 'c' value for the error term

	METHODS	GME: different 'c' VALUES of the fixed points					MLE
PARAMETERS		[-1S, 1S]	[-2S, 2S]	[-3S, 3S]	[-4S, 4S]	[-5S, 5S]	
B	***Bias***	-0,0067	-0,0107	-0,0107	-0,0117	-0,0139	0,1715
	MSE	0,0107	0,0114	0,011	0,0129	0,014	0,0394
C	***Bias***	-0,0144	-0,0143	-0,0029	-0,0098	-0,0111	0,0504
	MSE	0,0023	0,0021	0,0022	0,0019	0,0018	0,0034

Experiment 4. Likewise experiment 2, the simulation trial were repeated by shifting the support points of the error term to be equally spaced in the interval [-3 S_y, 0, 3 S_y]. The results of this experiment are in Table 6 suggested that we should fix three support values of the error term according to the three sigma rule.

Table 6. Fixing the number of support points for the error term

	METHODS	GME: different 'c' VALUES of the fixed points					MLE
PARAMETERS		3	4	5	6	7	
B	***Bias***	-0,0107	-0,0003	-0,0032	-0,0005	-0,0023	0,1715
	MSE	0,011	0,0169	0,0191	0,0168	0,0165	0,0394
C	***Bias***	-0,0029	-0,0138	-0,0092	-0,0127	-0,0161	0,0504
	MSE	0,0022	0,0021	0,0028	0,0022	0,0021	0,0034

5.2. Performance Analysis of the RSS technique

In this section, based on the main simulation conditions and the results of the last four experiments, we are comparing between two the estimation methods from the view of point of the sampling techniques. Two experiments were performed, the first one based on SRS and the other one based on RSS. Noting that, in the RSS scheme the ranking done based on the X variable. Other ranking schemes could be used such as Double RSS (Al-Saleh and Al-Kadiri, 2000), which is not of our interest in this paper, to be an alternative to the classical ranking scheme.

Experiment 5. Based on the results of experiment 1 to experiment 4, by selecting three support values for each of the parameters in the [-5, 5] interval and three support values the error term according to the three sigma rule, this experiment, under the simulation assumption outlined above, is conducted by increasing the sample size *n*, i.e., *n* = 20, 25, 30, 40 and 50. The results in Table 7 indicate that the GME is more accurate and more efficient than the MLE estimates for all sample sizes.

Table 7. Comparisons between GME and MLE using SRS

N	Method	C			B		
		Bias	MSE	eff	Bias	MSE	eff
20	GME	-0,0029	0,0022	1,54	-0,0107	0,0110	3,58
	MLE	0,0504	0,0034		0,1715	0,0394	
25	GME	-0,0150	0,0029	1,14	0,0043	0,0138	2,44
	MLE	0,0496	0,0033		0,1597	0,0337	
30	GME	-0,0085	0,0029	1,21	-0,0005	0,0096	4,11
	MLE	0,0510	0,0035		0,1726	0,0395	
40	GME	-0,0046	0,0018	1,67	-0,0006	0,0090	4,09
	MLE	0,0474	0,0030		0,1663	0,0368	
50	GME	-0,0013	0,0012	2,08	0,0007	0,0086	3,36
	MLE	0,0429	0,0025		0,1482	0,0289	

Experiment 6. The aim of this experiment is to improve the simulation results by using the simulated RSS. For the RSS sampling scheme we used a set size $r = 4$ or 5 and the number of cycles $m = 5,6,8$ and 10; to achieve the desired sample size. The results of this experiment are given in Table 8.

Table 8. Comparisons between GME and MLE using RSS

r	m	Method	C			B		
			Bias	MSE	Eff	Bias	MSE	eff
5	4	GME	0,0005	0,0056	1,76	0,0039	0,0059	8,55
		MLE	0,0532	0,0099		0,1938	0,0505	
5	5	GME	-0,0302	0,0053	1,85	0,0154	0,0144	3,55
		MLE	0,0543	0,0098		0,1953	0,0511	
6	5	GME	-0,0101	0,0025	1,64	-0,0003	0,0128	3,89
		MLE	0,0551	0,0041		0,1917	0,0499	
8	5	GME	-0,0179	0,0032	1,22	-0,0161	0,0107	3,78
		MLE	0,0506	0,0039		0,1726	0,0405	
10	5	GME	0,0078	0,0033	1,03	-0,0003	0,0084	6,12
		MLE	0,0535	0,0034		0,1929	0,0514	

We conclude by summarizing our findings:

1. In this study we used simulated ranked data with GME and MLE estimation methods to estimate the parameters of the Gompertz's model.
2. Under the simulation assumptions, GME based on both sampling techniques SRS and RSS gives a better estimate than MLE from the MSE point of view. The simulated efficiency score is more than 1 in both experiments for all parameters and under different sample sizes.

6. Conclusions and Discussion

To sum up, we consider fitting the Gompertz's model by two different estimation methods MLE and GME and based on two different sampling techniques: SRS and RSS. Despite the sampling technique, the simulation results demonstrate that the GME estimates are superior and often more efficient than MLE estimates in terms of MSE.

Table 9. Simulation results GME (SRS vs RSS)

N	Method	B			C		
		Bias	***MSE***	***Eff***	***Bias***	***MSE***	***eff***
20	***SRS***	-0,0107	0,0110	1,86	-0,0029	0,0022	0,39
	RSS	0,0039	0,0059		0,0005	0,0056	
25	***SRS***	0,0043	0,0138	0,96	-0,015	0,0029	0,55
	RSS	0,0154	0,0144		-0,0302	0,0053	
30	***SRS***	-0,0005	0,0096	0,75	-0,0085	0,0029	1,16
	RSS	-0,0003	0,0128		-0,0101	0,0025	
40	***SRS***	-0,0006	0,0090	0,84	-0,0046	0,0018	0,56
	RSS	-0,0161	0,0107		-0,0179	0,0032	
50	***SRS***	0,0007	0,0086	1,02	-0,0013	0,0012	0,36
	RSS	-0,0003	0,0084		0,0078	0,0033	

For all situations used in the simulation study considering different sample size, the GME estimates are more efficient and more accurate than the MLE estimates. However, there is no improvement in parameter estimation by MLE rely on RSS.

However, Table 9 shows the results obtained by comparing the SRS and RSS sampling schemes in the context of GME. We find out that they are comparable and no method is better than any other method, therefore, we can used GME as a robust estimation method to fit Gompertz's model whatever the sample technique is used.

Moreover it is widely shown in the literature that in empirical circumstances, the RSS can be employed to gain more information than SRS while keeping the cost of, or the time constraint on, the sampling about the same.

Consequently, the GME estimator based on any sampling scheme can be recommended for estimating the parameter of Gompertz's model as an alternative estimation method to the classical MLE.

Moreover, other statistical procedures and new methodologies in the context of RSS are proposed, like extreme ranked set sampling (ERSS), median ranked set sampling (MRSS) or L ranked set sampling (LRSS) by Al-Nasser (2007). It may be useful to investigate these different methods for the Gompertz's model, considering the GME estimator.

References

1. Al-Nasser A. D. **L Ranked Set Sampling: A Generalization Procedure for Robust Visual Sampling**, Communications in Statistics—Simulation and Computation, 36, 2007, pp. 33–43
2. Al-Saleh, M. F. and Al-Kadiri, M. **Double ranked set sampling,** Statistical Probability Letter, 48, 2000, pp. 205-212
3. Booth H. **Transforming Gompertz's Function for Fertility Analysis: The Development of a Standard for the Relational Gompertz Function**, Population Studies, Vol. 38, No. 3, Nov. 1984, pp. 495-506
4. Carey, J.R. **Insect biogeography**, Ann. Rev. Entomol. 46, 2001, pp. 79–110
5. Chen, Z., Bai, Z.D. and Sinha, B. K. **Ranked Set Sampling: Theory and Applications,** New York: Springer- Verlag, 2004

6. Ciavolino, E. **Modelling GME and PLS estimation methods for evaluating the Job Satisfaction in the Public Sector**, QMod'07, University of Helsingborg, Sweden, 18 – 21 June 2007, June 2007, Linköping University Electronic Press, ISSN number: 1650-3740
7. Ciavolino, E. and Dahlgaard, J. J. **Simultaneous Equation Model based on Generalized Maximum Entropy for studying the effect of the Management's Factors on the Enterprise Performances**, Journal of Applied Statistics, Vol. 36, Issue 7, 2009, pp. 801 – 815
8. Ciavolino, E. and Al-Nasser, A. D. **Comparing Generalized Maximum Entropy and Partial Least Squares methods for Structural Equation Models**, *Journal of Nonparametric Statistics*, Vol. 21, Issue 8, 2009, pp.1017 - 1036
10. Garg M. L., Rao B. R. and Redmond C. K. **Maximum-Likelihood Estimation of the Parameters of the Gompertz Survival Function**, Applied Statistics, Vol. 19, No. 2. 1970, pp. 152-159
11. Golan, A. **Information and Entropy Econometrics — A Review and Synthesis**, Foundations and Trends in Econometrics, 2, 1–2, 2008, pp. 1–145
12. Golan, A., Judge, G. and Karp, L. **A maximum entropy approach to estimation and inference in dynamic models or counting fish in the sea using maximum entropy**, Journal of Economic Dynamics and Control, 20, 2000, pp. 559-582
13. Golan, A., Karp, L. S. and Perloff, J. **Estimating Coke's and Pepsi's Price and Advertising Strategies**, Journal of Business & Economic Statistics, Vol. 18, No. 4, Oct. 2000, pp. 398-409
14. Golan, A., Perloff, J. M. and Shen, E. Z. **Estimating a Demand System with Nonnegativity Constraints: Mexican Meat Demand**, The Review of Economics and Statistics, Vol. 83, No. 3., Aug., 2001, pp. 541-550
15. Golan, A., George, G. and Miller, D. **Maximum Entropy Econometrics**, Wiley, London, 1996
16. Gompertz, B. **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 Royal Society A 115, 1825, pp. 513–580.
17. H. L. Rietz **On Certain Properties of Makeham's Laws of Mortality with Applications**, The American Mathematical Monthly, Vol. 28, No. 4, Apr., 1921, pp. 158-165.
18. Jaynes E.T. **Information Theory and Statistical Mechanics**, The Physical Review 106 (4), 620-630, May 15, 1957a
19. Jaynes E.T. **Information Theory and Statistics Methods II**, The Physical Review. 108, pp.171-90, 1957b
20. Jaynes, E. T. **Prior Probabilities**, IEEE Transactions On Systems Science and Cybernetics, vol. sec-4, no. 3, 1968, pp. 227-241
21. Mahajan V., Muller E. and Bass F. M. **New Product Diffusion Models in Marketing: A Review and Directions for Research**, Journal of Marketing, Vol. 54, No. 1, Jan. 1990, pp. 1-26
22. McIntyre, G. A. **A method of unbiased selective sampling using ranked sets**, Australian J. Agricultural Research, 3, 1952, pp. 385-390
23. Pukelsheim, F. **The Three Sigma Rule**, The American Statistician, 48(2), 1994, pp. 88–91
24. Ricklefs R. E., Scheuerlein A. **Biological Implications of the Weibull and Gompertz Models of Aging**, The Journals of Gerontology Series A: Biological Sciences and Medical Sciences 57:B69-B76, 2002
25. Schittkowski, K. **NLPQL: A FORTRAN subroutine solving constrained nonlinear programming problems**, Annals of Operations Research, 5, 1986, pp. 485–500

26. Shannon C. E. **A mathematical Theory of Communications**, Bell System Technical Journal, 27, 1948, pp. 379–423
27. Sinha, B. K., Sinha, B. K., and Purkayastha, S. **On some aspects of ranked set sampling for estimation of normal and exponential parameters**, Statist. Decisions 14, 1996, pp. 223–240
28. Willemse W. J. and Koppelaar H. **Knowledge Elicitation of Gompertz' Law of Mortality**, Scand. Actuarial J., 2000, 2, pp.168 -179
29. Wolfe D. A. **Ranked Set Sampling: an Approach to More Efficient Data Collection**, Statistical Science, Vol. 19, No. 4, 2004, pp. 636–643
30. Yamada S. and Osaki S. **Software Reliability Growth Modeling: Models and Applications**, IEEE Transactions on Software Engineering, vol. 11, no.12, Dec. 1985, pp. 1431-1437

[1] The estimates will not be in closed form and to get the final values, a numerical optimization technique (successive quadratic programming method) should be used to compute probabilities.