

Article

# Estimation and Prediction for Gompertz Distribution under General Progressive Censoring

Yuxuan Wang and Wenhao Gui \* 

Department of Mathematics, Beijing Jiaotong University, Beijing 100044, China; 18271204@bjtu.edu.cn

\* Correspondence: whgui@bjtu.edu.cn

**Abstract:** In this article, we discuss the estimation of the parameters for Gompertz distribution and prediction using general progressive Type-II censoring. Based on the Expectation–Maximization algorithm, we calculate the maximum likelihood estimates. Bayesian estimates are considered under different loss functions, which are symmetrical, asymmetrical and balanced, respectively. An approximate method—Tierney and Kadane—is used to derive the estimates. Besides, the Metropolis-Hastings (MH) algorithm is applied to get the Bayesian estimates as well. According to Fisher information matrix, we acquire asymptotic confidence intervals. Bootstrap intervals are also established. Furthermore, we build the highest posterior density intervals through the sample generated by the MH algorithm. Then, Bayesian predictive intervals and estimates for future samples are provided. Finally, for evaluating the quality of the approaches, a numerical simulation study is implemented. In addition, we analyze two real datasets.



**Citation:** Wang, Y.; Gui, W. Estimation and Prediction for Gompertz Distribution under General Progressive Censoring. *Symmetry* **2021**, *13*, 858. <https://doi.org/10.3390/sym13050858>

Academic Editors: Célia Nunes, Miguel Fonseca and Elsa Moreira

Received: 24 April 2021

Accepted: 7 May 2021

Published: 12 May 2021

**Publisher's Note:** MDPI stays neutral with regard to jurisdictional claims in published maps and institutional affiliations.



**Copyright:** © 2021 by the authors. Licensee MDPI, Basel, Switzerland. This article is an open access article distributed under the terms and conditions of the Creative Commons Attribution (CC BY) license (<https://creativecommons.org/licenses/by/4.0/>).

## 1. Introduction

Gompertz distribution has wide applications in describing human mortality, establishing actuarial tables and other fields. Historically, it was originally introduced by Gompertz (see Reference [1]). The probability density function (PDF) and cumulative distribution function (CDF) of the Gompertz distribution are defined as

$$f(x|\alpha, \beta) = \alpha \beta e^{\beta x} e^{-\alpha(e^{\beta x} - 1)}, \quad 0 < x < +\infty, \quad (1)$$

and

$$F(x|\alpha, \beta) = 1 - e^{-\alpha(e^{\beta x} - 1)}, \quad 0 < x < +\infty, \quad (2)$$

where the unknown parameters  $\alpha$  and  $\beta$  are positive.

The Gompertz distribution possesses a unimodal PDF; in addition to this, it also has an increasing hazard function. Many researchers have contributed to the properties of the Gompertz distribution. In recent years, Reference [2] studied the relations between other distributions and Gompertz distribution, for instance, the Type I extreme value and Weibull distributions. Reference [3] obtained the weighted and unweighted least squares estimations under censored and complete samples. Reference [4] calculated the maximum likelihood estimates (MLEs), and completed the establishment for exact confidence interval and joint confidence region base on progressive Type-II censoring. Reference [5] studied the statistical inferences for Gompertz distribution under generalized progressively hybrid censoring. They derived the MLEs by Newton's iteration method and used Markov chain Monte Carlo method to obtain Bayes estimates under generalized entropy and other loss functions. Bayesian predictions based on this censoring scheme were provided by one- and two-sample predictive approaches. Finally, they compared the proposed methods by simulation. Reference [6] obtained the MLEs and Bayesian estimates for the parameters

using progressive first-failure censoring, also the estimates of hazard rate and reliability functions of Gompertz distribution. Besides, approximate and exact confidence intervals were constructed, and the conjugate and discrete prior distributions for the parameters were proposed. Finally, a numerical example was reported. One may also refer to References [7,8] for extensions about Gompertz distribution.

In life tests and reliability analyses, censoring has attracted more and more attention due to time and cost savings. Several schemes of censoring are proposed in the literature, among which Type-I and Type-II censoring are the most widely used. The former allows an experiment to be ceased at a fixed time point and the number of observed failed units is random, while the latter asks the life testing to be stopped when a prescriptive number of units fail, and the duration of the experiment is random. Both the traditional Type-I and Type-II censoring methods have a limitation that surviving experimental units are only allowed to be withdrawn at the terminal point. In this regard, progressive type-II censoring has better practicability and flexibility because it allows removal of the surviving units after any failure occurs. However, sometimes there are also other cases in the test, such as the existence of the unobserved failures at the starting point of the test. This would result in more general censoring. In this article, following Reference [9], we concentrate on the general progressive Type-II censoring. Assume that a life testing contains  $n$  experimental units. The first  $r$  failures occur at the time points  $X_1, \dots, X_r$  respectively which are unobserved. When  $(r+1)$ -th failure is observed at the time point  $X_{r+1}$ , the surviving experimental units of size  $R_{r+1}$  are withdrawn, and so forth. When  $(r+i)$ -th failure takes place at the point of time  $X_{r+i}$  ( $i = 1, \dots, m-r$ ), the surviving experimental units of size  $R_{r+i}$  are removed at random. Eventually, when the  $m$ -th failed unit is observed at the time point  $X_m$ , remaining experimental units of size  $R_m = n - R_{r+1} - \dots - R_{m-1} - m$  are removed. Here  $R = (R_{r+1}, \dots, R_m)$  is prefixed and is cited as the censoring scheme, besides,  $X = (X_{r+1}, \dots, X_m)$  is known as general progressive censored data which denotes the observed failure time with size  $m-r$ .

Several scholars have discussed various lifetime distributions using the general progressive censored data. Among others, Reference [10] derived both classical and Bayesian estimates using general progressive censored data obtained randomly from exponential distribution. Reference [11] applied general progressive censored sample to discuss Bayesian estimates for the two parameters in inverse Weibull distribution and prediction problems with the priors with gamma distribution on the scale parameter and a log-concave density on the shape parameter. Reference [12] obtained Bayesian prediction estimates for the future sample from Weibull distribution using asymmetric and symmetric loss functions under general progressive censoring. Other studies can be found in References [13–15], and so forth.

In this article, using general progressive censoring, we discuss the estimation and prediction problems on Gompertz distribution.

This paper proceeds as follows: First, we calculate the MLEs in Section 2 by the Expectation-Maximization (EM) method and acquire the Fisher information matrix. Besides, in the same section we also derive the bootstrap intervals. In Section 3, we discuss the Bayesian estimates with different loss functions. An approximate method, Tierney and Kadane (TK), is proposed to calculate these estimates. Furthermore, we apply the MH algorithm to obtain Bayesian estimations and establish highest posterior density (HPD) intervals under the sample generated by the MH algorithm. In Section 4, Bayesian point prediction and interval prediction estimates for future samples are provided. A numerical simulation is executed in Section 5 to evaluate the quality of these approaches, in addition, we also analyze two real datasets. Finally, conclusions are arranged in the last Section 6.

## 2. Maximum Likelihood Estimation

Let  $R = (R_{r+1}, \dots, R_m)$  be the censoring scheme, under which  $X = (X_{r+1}, X_{r+2}, \dots, X_m)$  denotes the corresponding general progressive censored sample drawn from Gompertz distribution. Then, the likelihood function is derived as the following expression:

$$l(\alpha, \beta | \underline{x}) = Q[F(x_{r+1} | \alpha, \beta)]^r \prod_{i=r+1}^m [1 - F(x_{r+1} | \alpha, \beta)]^{R_i} f(x_{r+1} | \alpha, \beta), \quad (3)$$

where  $Q = \binom{n}{r}(n-r) \prod_{j=r+2}^m (n - \sum_{i=r+1}^{j-1} R_i + 1 - j)$ , and  $\underline{x} = (x_{r+1}, \dots, x_m)$  denotes an observed value of  $X = (X_{r+1}, \dots, X_m)$ . In addition,  $\binom{n}{r}$  is the binomial coefficient, that is  $\frac{n!}{(n-r)!r!}$ .

Substituting (1) and (2), (3) can be written in the following form:

$$l(\alpha, \beta | \underline{x}) = Q\alpha^{m-r}\beta^{m-r}[1 - e^{-\alpha(e^{\beta x_{r+1}} - 1)}]^r \prod_{i=r+1}^m e^{\beta x_i} e^{-\alpha(R_i + 1)(e^{\beta x_i} - 1)}. \quad (4)$$

### 2.1. Point Estimation with EM Algorithm

A classical method for obtaining MLE is the Newton–Raphson method, which requires the second-order partial derivatives of the log-likelihood function and the derivatives are usually complicated in the case of censoring. Therefore, it is necessary to seek other methods. Following Reference [16], we use the EM algorithm to derive the MLEs. This algorithm is powerful for handling incomplete data problems because only the pseudo-log likelihood function of complete data needs to be maximized. It is an iterative method by using current estimation of the parameters to expect the log-likelihood function filled with censored data which is called E-step and maximize it to get the next estimation which is called M-step.

We use  $Z = (Z_r, Z_{r+1}, \dots, Z_m)$  to represent the censored sample, where  $Z_r$  is a  $1 \times r$  vector  $Z_r = (Z_{r1}, \dots, Z_{rr})$  of the first  $r$  unobserved failures, and  $Z_i, i = r+1, \dots, m$  denotes a  $1 \times R_i$  vector  $Z_i = (Z_{i1}, \dots, Z_{iR_i})$  of the censored data after  $X_i$  failed. The observed and complete sample are denoted by  $X = (X_{r+1}, \dots, X_m)$  and  $K$  respectively, then  $K = (X, Z)$ . Let  $\underline{z}_0 = (z_{r1}, \dots, z_{rr})$ ,  $\underline{z}_i = (z_{i1}, \dots, z_{iR_i})$  ( $i = r+1, \dots, m$ ) and  $\underline{x} = (x_{r+1}, \dots, x_m)$  represent the corresponding observations. Under the complete data, we can express the log-likelihood function by

$$\begin{aligned} L_c(\alpha, \beta, K) = & n \ln(\alpha\beta) + \beta \sum_{i=r+1}^m x_i - \alpha \sum_{i=r+1}^m (e^{\beta x_i} - 1) \\ & + \beta \sum_{i=r+1}^m \sum_{k=1}^{R_i} z_{ik} - \alpha \sum_{i=r+1}^m \sum_{k=1}^{R_i} (e^{\beta z_{ik}} - 1) - \alpha \sum_{k=1}^r (e^{\beta z_{rk}} - 1) + \beta \sum_{k=1}^r z_{rk}. \end{aligned} \quad (5)$$

- E-step

To conduct the E-step smoothly, first we compute the expectation of (5), the pseudo-log likelihood function is then expressed as

$$\begin{aligned} L_E(\alpha, \beta; \underline{x}) = & n \ln(\alpha\beta) + \beta \sum_{i=r+1}^m x_i - \alpha \sum_{i=r+1}^m (e^{\beta x_i} - 1) \\ & + \beta \sum_{k=1}^r E(z_{rk} | z_{rk} < x_{r+1}) + \beta \sum_{i=r+1}^m \sum_{k=1}^{R_i} E(z_{ik} | z_{ik} > x_i) \\ & - \alpha \sum_{k=1}^r E((e^{\beta z_{rk}} - 1) | z_{rk} < x_{r+1}) - \alpha \sum_{i=r+1}^m \sum_{k=1}^{R_i} E((e^{\beta z_{ik}} - 1) | z_{ik} > x_i), \end{aligned} \quad (6)$$

where

$$\begin{aligned} E(z_{ik}|z_{ik} > x_i) &= \frac{1}{1 - F(x_i|\alpha, \beta)} \int_{x_i}^{+\infty} \alpha \beta e^{\beta t} e^{-\alpha(e^{\beta t} - 1)} t dt \\ &= \frac{\alpha}{\beta e^{-\alpha(e^{\beta x_i} - 1)}} \int_{e^{\beta x_i} - 1}^{+\infty} \ln(t + 1) e^{-\alpha t} dt \\ &= E_1(x_i, \alpha, \beta), \end{aligned} \quad (7)$$

$$\begin{aligned} E((e^{\beta z_{ik}} - 1)|z_{ik} > x_i) &= \frac{1}{1 - F(x_i|\alpha, \beta)} \int_{x_i}^{+\infty} \alpha \beta e^{\beta t} e^{-\alpha(e^{\beta t} - 1)} (e^{\beta t} - 1) dt \\ &= e^{\beta x_i} - 1 + \frac{1}{\alpha} \\ &= E_2(x_i, \alpha, \beta), \end{aligned} \quad (8)$$

$$\begin{aligned} E(z_{rk}|z_{rk} < x_{r+1}) &= \frac{1}{F(x_{r+1}|\alpha, \beta)} \int_0^{x_{r+1}} \alpha \beta e^{\beta t} e^{-\alpha(e^{\beta t} - 1)} t dt \\ &= \frac{\alpha}{\beta F(x_{r+1}|\alpha, \beta)} \int_0^{e^{\beta x_{r+1}} - 1} e^{-\alpha t} \ln(t + 1) dt \\ &= E_3(x_{r+1}, \alpha, \beta), \end{aligned} \quad (9)$$

and

$$\begin{aligned} E((e^{\beta z_{rk}} - 1)|z_{rk} < x_{r+1}) &= \frac{1}{F(x_{r+1}|\alpha, \beta)} \int_0^{x_{r+1}} \alpha \beta e^{\beta t} e^{-\alpha(e^{\beta t} - 1)} (e^{\beta t} - 1) dt \\ &= \frac{1 - e^{\beta x_{r+1}} e^{-\alpha(e^{\beta x_{r+1}} - 1)}}{F(x_{r+1}|\alpha, \beta)} - 1 + \frac{1}{\alpha} \\ &= E_4(x_{r+1}, \alpha, \beta). \end{aligned} \quad (10)$$

#### • M-step

Suppose that the  $s$ -th estimate of  $(\alpha, \beta)$  is represented by  $(\alpha_{(s)}^*, \beta_{(s)}^*)$  then the M-step aims to maximize (6) by substituting  $\alpha_{(s)}^*$  and  $\beta_{(s)}^*$  into  $E_1, E_2, E_3$  and  $E_4$ , and derive the  $(s+1)$ -th estimate. Therefore, the next task is to maximize the function

$$\begin{aligned} L_M(\alpha, \beta; x) &= n \ln(\alpha \beta) + \beta \sum_{i=r+1}^m x_i - \alpha \sum_{i=r+1}^m (e^{\beta x_i} - 1) + \beta \sum_{i=r+1}^m R_i E_1^* \\ &\quad - \alpha \sum_{i=r+1}^m R_i E_2^* + r \beta E_3^* - r \alpha E_4^*, \end{aligned} \quad (11)$$

where  $E_1^*, E_2^*, E_3^*, E_4^*$  are  $E_1(x_i, \alpha_{(s)}^*, \beta_{(s)}^*), E_2(x_i, \alpha_{(s)}^*, \beta_{(s)}^*), E_3(x_{r+1}, \alpha_{(s)}^*, \beta_{(s)}^*), E_4(x_{r+1}, \alpha_{(s)}^*, \beta_{(s)}^*)$ , respectively. The corresponding likelihood equations are

$$\frac{\partial L_M}{\partial \alpha} = \frac{n}{\alpha} - \sum_{i=r+1}^m (e^{\beta x_i} - 1 + R_i E_2^*) - r E_4^* = 0, \quad (12)$$

and

$$\frac{\partial L_M}{\partial \beta} = \frac{n}{\beta} + \sum_{i=r+1}^m (x_i - \alpha x_i e^{\beta x_i} + R_i E_1^*) + r E_3^* = 0. \quad (13)$$

For the reason that it is infeasible to compute (12) and (13) analytically, we use a numerical technique to obtain  $\alpha_{(s+1)}^*$  and  $\beta_{(s+1)}^*$ . From (12), the estimate of  $\alpha$  can be described as the following function of  $\beta$ :

$$\hat{\alpha} = \frac{n}{\sum_{i=r+1}^m (e^{\beta x_i} - 1 + R_i E_2^*) + r E_4^*}. \quad (14)$$

By replacing  $\alpha$  with  $\hat{\alpha}$ , Equation (13) can be transformed into the equivalent form  $\beta = FF(\beta)$ , where

$$FF(\beta) = \frac{n}{\sum_{i=r+1}^m (\hat{\alpha} x_i e^{\beta x_i} - x_i + R_i E_1^*) - r E_3^*}. \quad (15)$$

Then,  $\beta_{(s+1)}^*$  can be acquired using the fixed-point iterative procedure:

$$\beta_{j+1} = FF(\beta_j). \quad (16)$$

When  $|\beta_{j+1} - \beta_j|$  is smaller than a given tolerance limit, the iteration stops. Once we get  $\beta_{(s+1)}^*$ ,  $\alpha_{(s+1)}^*$  can be computed as  $\alpha_{(s+1)}^* = \hat{\alpha}(\beta_{(s+1)}^*)$  easily from (14). Repeat E-step and M-step till this program converges. Then we get the MLEs for  $\alpha$  and  $\beta$ .

## 2.2. Asymptotic Confidence Interval

Now, we acquire the Fisher information matrix and establish  $100(1 - \gamma)\%$  asymptotic confidence intervals (ACIs).

The observed information can be extracted by means of a program proposed by Reference [17] when using EM method in handling incomplete sample problems to derive MLEs. Let  $\theta$  be the unknown parameter  $(\alpha, \beta)$ .  $I_K(\theta)$  and  $I_X(\theta)$  denote complete information and observed information respectively. Furthermore, missing information is represented as  $I_{K|X}(\theta)$ . The main concept of this program can be described as the principle of missing information

$$I_X(\theta) = I_K(\theta) - I_{K|X}(\theta), \quad (17)$$

where  $I_K(\theta)$  can be derived by

$$I_K(\theta) = -E \left[ \frac{\partial^2 L_c(\theta, K)}{\partial \theta^2} \right] = \begin{bmatrix} A_{11} & A_{12} \\ A_{21} & A_{22} \end{bmatrix}. \quad (18)$$

$I_{W|X}$  is the expected information of the distribution of  $Z = (Z_r, Z_{r+1}, \dots, Z_m)$  given  $X = (X_{r+1}, \dots, X_m)$ . According to Reference [18], given the observed sample of general progressive Type-II censoring, we obtain the distribution of  $Z$  as

$$f(Z|X, \theta) = \prod_{k=1}^r \frac{f(z_{rk}|\theta)}{F(x_{r+1}|\theta)} \prod_{i=r+1}^m \prod_{k=1}^{R_i} \frac{f(z_{ik}|\theta)}{1 - F(x_i|\theta)}, \quad (19)$$

then the  $I_{K|X}$  is

$$I_{K|X}(\theta) = -E \left[ \frac{\partial^2 \ln f(Z|X, \theta)}{\partial \theta^2} \right]. \quad (20)$$

Let  $f_{Z_0}(z|x_{r+1}, \theta) = f(z|\theta)/F(x_{r+1}|\theta)$ ,  $f_{Z_i}(z_i|x_i, \theta) = f(z_i|\theta)/[1 - F(x_i|\theta)]$ . Following Reference [18], we know that given the observed sample  $(X_{r+1}, \dots, X_m) = (x_{r+1}, \dots, x_m)$ , the components of  $Z_r$  are independent of each other and have the PDF  $f_{Z_0}(z_{rk}|x_{r+1}, \theta)$ ,  $k = 1, \dots, r$ . Similarly, the components of  $Z_i$ ,  $i = r+1, \dots, m$  are independent of each other and have the PDF  $f_{Z_i}(z_{ik}|x_i, \theta)$ ,  $k = 1, \dots, R_i$ . Therefore, the  $I_{K|X}$  can be restated as

$$I_{K|X}(\theta) = rI_{K|X}^*(\theta) + \sum_{i=r+1}^m R_i I_{K|X}^{(i)}(\theta), \quad (21)$$

where

$$I_{K|X}^*(\theta) = -E \left[ \frac{\partial^2 \ln f_{Z_0}(z|x_{r+1}, \theta)}{\partial \theta^2} \right] = \begin{bmatrix} B_{11} & B_{12} \\ B_{21} & B_{22} \end{bmatrix}, \quad (22)$$

and

$$I_{K|X}^{(i)}(\theta) = -E \left[ \frac{\partial^2 \ln f_{Z_i}(z_i|x_i, \theta)}{\partial \theta^2} \right] = \begin{bmatrix} C_{11} & C_{12} \\ C_{21} & C_{22} \end{bmatrix}. \quad (23)$$

Now we can figure out the elements of the above matrices as follows:

$$\begin{aligned} A_{11} &= n/\alpha^2, \quad A_{21} = A_{12} = \frac{n\alpha}{\beta} \int_0^{+\infty} e^{-\alpha t} (t+1) \ln(t+1) dt, \\ A_{22} &= \frac{n}{\beta^2} + \frac{n\alpha}{\beta^2} \int_0^{+\infty} e^{-\alpha t} (t+1) [\ln(t+1)]^2 dt, \\ B_{11} &= -\frac{1}{\alpha^2} + h_1, \quad B_{12} = B_{21} = -\frac{\alpha}{\beta F(x_{r+1}|\theta)} \int_0^{e^{\beta x_{r+1}}-1} e^{-\alpha t} (t+1) \ln(t+1) dt + h_2, \\ B_{22} &= -\frac{1}{\beta^2} - \frac{\alpha^2}{\beta^2 F(x_{r+1}|\theta)} \int_0^{e^{\beta x_{r+1}}-1} e^{-\alpha t} (t+1) [\ln(t+1)]^2 dt + h_3, \\ C_{11} &= -\frac{1}{\alpha^2}, \quad C_{12} = C_{21} = -\frac{\alpha}{\beta(1-F(x_i|\theta))} \int_{e^{\beta x_i}-1}^{+\infty} e^{-\alpha t} (t+1) \ln(t+1) dt + x_i e^{\beta x_i}, \\ C_{22} &= -\frac{\alpha^2}{\beta^2(1-F(x_i|\theta))} \int_{e^{\beta x_i}-1}^{+\infty} e^{-\alpha t} (t+1) [\ln(t+1)]^2 dt - \frac{1}{\beta^2} + \alpha x_i^2 e^{\beta x_i}, \end{aligned}$$

where

$$\begin{aligned} h_1 &= \frac{(e^{\beta x_{r+1}} - 1)^2 e^{-\alpha(e^{\beta x_{r+1}}-1)}}{[F(x_{r+1}|\theta)]^2}, \\ h_2 &= -\frac{x_{r+1} e^{\beta x_{r+1}} e^{-\alpha(e^{\beta x_{r+1}}-1)} [1 - \alpha(e^{\beta x_{r+1}} - 1) - e^{-\alpha(e^{\beta x_{r+1}}-1)}]}{[F(x_{r+1}|\theta)]^2}, \end{aligned}$$

and

$$h_3 = -\frac{\alpha x_{r+1}^2 e^{\beta x_{r+1}} [1 - \alpha e^{\beta x_{r+1}} - e^{-\alpha(e^{\beta x_{r+1}}-1)}]}{[F(x_{r+1}|\theta)]^2}.$$

Further, using asymptotic normality of MLE  $\hat{\theta} = (\hat{\alpha}, \hat{\beta})$ ,  $\hat{\theta} \sim N(\hat{\theta}, I^{-1}(\hat{\theta}))$ , the  $100(1 - \gamma)\%$  ACIs for the two unknown parameters are obtained as

$$\left( \hat{\alpha} - \eta_{\frac{\gamma}{2}} \sqrt{Var(\hat{\alpha})}, \hat{\alpha} + \eta_{\frac{\gamma}{2}} \sqrt{Var(\hat{\alpha})} \right) \text{ and } \left( \hat{\beta} - \eta_{\frac{\gamma}{2}} \sqrt{Var(\hat{\beta})}, \hat{\beta} + \eta_{\frac{\gamma}{2}} \sqrt{Var(\hat{\beta})} \right), \quad (24)$$

where  $\eta_{\frac{\gamma}{2}}$  represents the upper  $\frac{\gamma}{2}$ -th quantile for standard normal distribution,  $Var(\hat{\alpha})$  and  $Var(\hat{\beta})$  denote the principal diagonal elements of  $I^{-1}(\hat{\theta})$  respectively.

### 2.3. Bootstrap Confidence Interval

As is widely known, asymptotic confidence interval on the basis of MLE requires a large sample to support its accuracy but, in many practical cases, the sample size tends to not be enough. Reference [19] proposed the bootstrap method to construct the confidence interval (CI), which is more suitable for small sample. In this part, the parametric bootstrap method is employed in the establishment of percentile bootstrap (bootstrap-p) and bootstrap-t CIs for a parameter  $\lambda$  (here  $\lambda$  is  $\alpha$  or  $\beta$ ). Interested readers may refer to References [20,21] for more information about bootstrap and Reference [22] for the algorithm of generating the general progressive type-II censored sample.

#### Parametric bootstrap-p

- (1) Calculate the MLEs  $\hat{\alpha}_0$  and  $\hat{\beta}_0$  based on the existing general progressive censored data and censoring scheme  $R = (R_{r+1}, \dots, R_m)$ .
- (2) Generate  $B_m$  from Beta( $n - r, r + 1$ ).
- (3) Generate independent  $U_{r+k}$  from Uniform(0, 1),  $k = 1, 2, \dots, m - r - 1$ .
- (4) Set  $B_{r+k} = U_{r+k}^{\frac{1}{\xi_{r+k}}}$ , where  $\xi_{r+k} = k + \sum_{i=m-k+1}^m R_i$ ,  $k = 1, 2, \dots, m - r - 1$ .
- (5) Set  $Z_{r+k} = 1 - B_{m-k+1} B_{m-k+2} \dots B_m$ ,  $k = 1, 2, \dots, m - r$ .
- (6) Set  $X_k = F^{-1}(Z_k)$ ,  $k = r + 1, \dots, m$ , and  $F(\cdot)$  represents the CDF of Gompertz distribution with parameters  $\hat{\alpha}_0$  and  $\hat{\beta}_0$ . Then the  $X_k$ ,  $k = r + 1, \dots, m$  are the general progressive censored sample (also bootstrap sample).
- (7) Compute the MLEs  $\hat{\alpha}^*$  and  $\hat{\beta}^*$  using the updated bootstrap sample.
- (8) Repeat steps (2)–(7)  $D$  times. Acquire the estimates:  $(\hat{\alpha}_1^*, \hat{\alpha}_2^*, \dots, \hat{\alpha}_D^*), (\hat{\beta}_1^*, \hat{\beta}_2^*, \dots, \hat{\beta}_D^*)$ .
- (9) Set  $\hat{F}_\lambda(x) = P(\hat{\lambda}^* \leq x)$  as the CDF for  $\hat{\lambda}^*$ . For a given value of  $x$ , define  $\hat{\lambda}_p(x) = \hat{F}_\lambda^{-1}(x)$ . The  $100(1 - \gamma)\%$  bootstrap-p CI of the parameter  $\lambda$  is obtained as

$$\left( \hat{\lambda}_p\left(\frac{\gamma}{2}\right), \hat{\lambda}_p\left(1 - \frac{\gamma}{2}\right) \right).$$

#### Parametric bootstrap-t

- (1)–(7) The same as the bootstrap-p above.
- (8) Obtain the statistics  $T_\lambda^*$  that

$$T_\lambda^* = \frac{\hat{\lambda}^* - \hat{\lambda}}{\sqrt{\widehat{Var}(\hat{\lambda}^*)}}.$$

- (9) Repeat steps (2)–(8)  $D$  times.
- (10) Set  $\hat{F}_{T_\lambda}(x) = P(T_\lambda^* \leq x)$  as the CDF for  $T_\lambda^*$ . For a given value of  $x$ , define  $\hat{\lambda}_t(x) = \hat{\lambda} + \sqrt{\widehat{Var}(\hat{\lambda}^*)} \hat{F}_{T_\lambda}^{-1}(x)$ . The  $100(1 - \gamma)\%$  bootstrap-t CI for the parameter  $\lambda$  is given by

$$\left( \hat{\lambda}_t\left(\frac{\gamma}{2}\right), \hat{\lambda}_t\left(1 - \frac{\gamma}{2}\right) \right).$$

### 3. Bayesian Estimation

Bayesian statistics are different from traditional statistics in that they allow the incorporation of subjective prior knowledge about life parameters into the inferential procedure in reliability analysis. Therefore, for the same quality of inferences, Bayesian methods tend to require fewer sample data than traditional statistical methods do. This makes it extremely important in expensive life tests.

We investigate the Bayesian estimates in this section. Suppose that  $\alpha$  and  $\beta$  independently have gamma prior distributions with the parameters  $(a, b)$  and  $(c, d)$ . Afterwards, we can obtain their joint prior distribution, that is

$$\pi(\alpha, \beta) = \frac{b^a d^c}{\Gamma(a)\Gamma(c)} \alpha^{a-1} e^{-b\alpha} \beta^{c-1} e^{-d\beta}, \quad 0 < \alpha, \beta < +\infty, \quad (25)$$

where the positive constants  $a, b, c$  and  $d$  are hyperparameters. Let  $\underline{x} = (x_{r+1}, \dots, x_m)$  be an observed value of  $X = (X_{r+1}, \dots, X_m)$ . Based on the joint prior distribution and likelihood function, the joint posterior function is

$$\begin{aligned} \pi(\alpha, \beta | \underline{x}) &= \frac{\pi(\alpha, \beta) l(\alpha, \beta | \underline{x})}{\int_0^{+\infty} \int_0^{+\infty} l(\alpha, \beta | \underline{x}) \pi(\alpha, \beta) d\alpha d\beta} \\ &\propto \alpha^{m+a-r-1} \beta^{m+c-r-1} e^{-(b\alpha+d\beta)} [1 - e^{-\alpha(e^{\beta x_{r+1}} - 1)}]^{r-1} \prod_{i=r+1}^m e^{\beta x_i} e^{-\alpha(R_i+1)(e^{\beta x_i} - 1)}. \end{aligned} \quad (26)$$

It is clear that (26) is analytically tricky. Furthermore, the Bayesian estimation of a function with  $\alpha$  and  $\beta$  is also intractable because it is related to a ratio of two integrals. For solving the corresponding ratio of two integrals, some approximate approaches have been presented in the literature. Among them, the TK method was proposed by Reference [23] to obtain the approximate posterior expectations. Besides, the MH algorithm is a simulation method with wide applications in sampling from posterior density function. In this article, we use the TK method and the MH algorithm to derive approximate explicit forms for the Bayesian estimates.

### 3.1. Loss Functions

In Bayesian statistics, the selection of loss function is a fundamental step. There are many symmetric loss functions, among which squared error loss (SEL) function is well-known for its good mathematical properties. Let  $\delta$  be a Bayesian estimate of  $\theta$ . The form of SEL function is

$$L_1(\delta, \theta) = (\delta - \theta)^2, \quad (27)$$

then under SEL function the Bayesian estimate for  $\theta$  can be computed by  $\delta_S = E^{(\theta|X)}(\theta|X)$ .

However, in many practical situations, overestimation and underestimation result in different losses, and the consequence is likely to be quite serious if one uses symmetric loss function indiscriminately. In these cases, asymmetrical loss functions are considered to be more suitable. In the literature, many different asymmetric loss functions were used. Among them, LINEX is dominant, and this loss function can be expressed as

$$L_2(\delta, \theta) = \zeta(e^{h(\delta-\theta)} - h(\delta-\theta) - 1), \quad h \neq 0, \zeta > 0. \quad (28)$$

Without loss of generality, here, we take  $\zeta = 1$ . Thus Bayesian estimate for  $\theta$  is given by the expression  $\delta_L = -\frac{1}{h} \ln\{E^{(\theta|X)}(e^{-h\theta}|X)\}$ .

Later, Reference [24] proposed a balanced loss function that has a more generalized form

$$L_3(\delta, \theta) = \sigma \rho(\delta, \delta_0) + (1 - \sigma) \rho(\delta, \theta), \quad 0 \leq \sigma \leq 1, \quad (29)$$

where  $\delta_0$  is a known estimate of  $\theta$  such as the MLE, and  $\rho$  is a loss function selected arbitrarily. By choosing  $\rho$  as SEL function given by (27), the (29) is transformed into balanced squared error loss (BSEL) function. According to BSEL we can give the Bayesian estimate by  $\delta_{BS} = (1 - \sigma)E^{(\theta|X)}(\theta|X) + \sigma\delta_0$ .

It can be clearly seen that the balanced loss function is more general since it includes special cases of the MLE, symmetric and asymmetric loss functions. For instance, by setting  $\delta_0$  to be the MLE of parameter, under BSEL function, the Bayesian estimate is exactly equal to MLE if  $\sigma = 1$ , and it is simplified into the Bayesian estimate under SEL function when  $\sigma = 0$ . Similarly, if  $\rho$  is chosen as LINEX loss function given by (28),  $L_3(\cdot)$  is called BLINEX

function. When  $\sigma = 1$  and  $\sigma = 0$ , the Bayesian estimates under the BLINEX loss function correspondingly reduce to MLE and the case of LINEX loss function.

In this article, we derive Bayesian estimates under SEL, BSEL functions and LINEX loss functions, respectively. Next, the TK method is suggested to deal with the ratio of the integrals problem on posterior expectation estimation.

### 3.2. TK Method

We assume that  $u(\alpha, \beta)$  denotes an arbitrary function of  $(\alpha, \beta)$ . Following Reference [23], the posterior expectation for  $u(\alpha, \beta)$  is written as

$$E(u(\alpha, \beta)) = \frac{\int_0^{+\infty} \int_0^{+\infty} u(\alpha, \beta) \pi(\alpha, \beta) e^{L(\alpha, \beta|x)} d\alpha d\beta}{\int_0^{+\infty} \int_0^{+\infty} e^{L(\alpha, \beta|x)} \pi(\alpha, \beta) d\alpha d\beta}, \quad (30)$$

where  $\pi(\alpha, \beta)$  denotes the prior density, and  $L(\alpha, \beta|x)$  represents the logarithm of (4). We set:

$$\varphi(\alpha, \beta) = \frac{\ln \pi(\alpha, \beta) + L(\alpha, \beta|x)}{n} \quad \text{and} \quad \varphi_u^*(\alpha, \beta) = \frac{\ln u(\alpha, \beta)}{n} + \varphi(\alpha, \beta). \quad (31)$$

Maximizing  $\varphi(\alpha, \beta)$  and  $\varphi_u^*(\alpha, \beta)$  individually, we derive  $(\hat{\alpha}_1, \hat{\beta}_1)$  and  $(\hat{\alpha}_u, \hat{\beta}_u)$ . Then, the approximate posterior expectation of  $u(\alpha, \beta)$  by applying the TK method is obtained as

$$\hat{E}(u(\alpha, \beta)) = \sqrt{\frac{|\Sigma_u^*|}{|\Sigma|}} \exp\{n[\varphi_u^*(\hat{\alpha}_u, \hat{\beta}_u) - \varphi(\hat{\alpha}_1, \hat{\beta}_1)]\}, \quad (32)$$

where  $|\Sigma|$  and  $|\Sigma_u^*|$  represent the corresponding determinants of negative inverse Hessian matrix of  $\varphi(\alpha, \beta)$  and  $\varphi_u^*(\alpha, \beta)$ . Next, ignoring the constant term, we note that

$$\begin{aligned} \varphi(\alpha, \beta) = & \frac{1}{n} \left[ (m-r) \ln(\alpha\beta) + \beta \sum_{i=r+1}^m x_i + r \ln[1 - e^{-\alpha(e^{\beta x_{r+1}} - 1)}] \right. \\ & \left. - \alpha \sum_{i=r+1}^m (R_i + 1)(e^{\beta x_i} - 1) - b\alpha + (a-1) \ln \alpha - d\beta + (c-1) \ln \beta \right]. \end{aligned} \quad (33)$$

Now, we compute the partial derivatives of  $\varphi$ :

$$\frac{\partial \varphi}{\partial \alpha} = \frac{1}{n} \left[ \frac{m-r}{\alpha} - \sum_{i=r+1}^m (e^{\beta x_i} - 1)(R_i + 1) - b + \frac{a-1}{\alpha} + \frac{r e^{-\alpha(e^{\beta x_{r+1}} - 1)} (e^{\beta x_{r+1}} - 1)}{F(x_{r+1}|\alpha, \beta)} \right], \quad (34)$$

and

$$\frac{\partial \varphi}{\partial \beta} = \frac{1}{n} \left[ \frac{m-r}{\beta} + \sum_{i=r+1}^m [x_i - \alpha(R_i + 1)e^{\beta x_i} x_i] + \frac{c-1}{\beta} - d + \frac{\alpha x_{r+1} e^{\beta x_{r+1}} e^{-\alpha(e^{\beta x_{r+1}} - 1)}}{F(x_{r+1}|\alpha, \beta)} \right]. \quad (35)$$

Similarly, the second derivatives can be derived as

$$\frac{\partial^2 \varphi}{\partial \alpha^2} = \frac{1}{n} \left[ -\frac{m+a-r-1}{\alpha^2} - \frac{r(e^{\beta x_{r+1}} - 1)^2 e^{-\alpha(e^{\beta x_{r+1}} - 1)}}{F^2(x_{r+1}|\alpha, \beta)} \right], \quad (36)$$

$$\begin{aligned} \frac{\partial^2 \varphi}{\partial \beta^2} = & \frac{1}{n} \left[ -\frac{m+c-r-1}{\beta^2} - \alpha \sum_{i=r+1}^m (R_i + 1) e^{\beta x_i} x_i^2 \right. \\ & \left. + \frac{r \alpha x_{r+1}^2 e^{\beta x_{r+1}} e^{-\alpha(e^{\beta x_{r+1}}-1)} [1 - \alpha e^{\beta x_{r+1}} - e^{-\alpha(e^{\beta x_{r+1}}-1)}]}{F^2(x_{r+1} | \alpha, \beta)} \right], \end{aligned} \quad (37)$$

and

$$\begin{aligned} \frac{\partial^2 \varphi}{\partial \beta \partial \alpha} = \frac{\partial^2 \varphi}{\partial \alpha \partial \beta} = & \frac{1}{n} \left[ - \sum_{i=r+1}^m (R_i + 1) e^{\beta x_i} x_i \right. \\ & \left. + \frac{r x_{r+1} e^{\beta x_{r+1}} e^{-\alpha(e^{\beta x_{r+1}}-1)} [1 - \alpha(e^{\beta x_{r+1}}-1) - e^{-\alpha(e^{\beta x_{r+1}}-1)}]}{F^2(x_{r+1} | \alpha, \beta)} \right]. \end{aligned} \quad (38)$$

Through (36)–(38),  $|\Sigma|$  is obtained as

$$|\Sigma| = \left[ \frac{\partial^2 \varphi}{\partial \alpha^2} \frac{\partial^2 \varphi}{\partial \beta^2} - \frac{\partial^2 \varphi}{\partial \beta \partial \alpha} \frac{\partial^2 \varphi}{\partial \alpha \partial \beta} \right]_{\alpha=\hat{\alpha}_1, \beta=\hat{\beta}_1}^{-1}. \quad (39)$$

For  $|\Sigma^*|$ , we first compute that

$$\frac{\partial \varphi^*}{\partial \alpha} = \frac{\partial \varphi}{\partial \alpha} + \frac{u_\alpha}{nu(\alpha, \beta)}, \quad (40)$$

$$\frac{\partial \varphi^*}{\partial \beta} = \frac{\partial \varphi}{\partial \beta} + \frac{u_\beta}{nu(\alpha, \beta)}, \quad (41)$$

$$\frac{\partial^2 \varphi^*}{\partial \alpha^2} = \frac{\partial^2 \varphi}{\partial \alpha^2} + \frac{1}{n} \left[ \frac{u(\alpha, \beta) u_{\alpha\alpha} - (u_\alpha)^2}{[u(\alpha, \beta)]^2} \right], \quad (42)$$

$$\frac{\partial^2 \varphi^*}{\partial \beta^2} = \frac{\partial^2 \varphi}{\partial \beta^2} + \frac{1}{n} \left[ \frac{u(\alpha, \beta) u_{\beta\beta} - (u_\beta)^2}{[u(\alpha, \beta)]^2} \right], \quad (43)$$

and

$$\frac{\partial^2 \varphi^*}{\partial \beta \partial \alpha} = \frac{\partial^2 \varphi^*}{\partial \alpha \partial \beta} = \frac{\partial^2 \varphi}{\partial \alpha \partial \beta} + \frac{1}{n} \left[ \frac{u(\alpha, \beta) u_{\alpha\beta} - u_\alpha u_\beta}{[u(\alpha, \beta)]^2} \right]. \quad (44)$$

As a result,  $|\Sigma^*|$  is

$$|\Sigma^*| = \left[ \frac{\partial^2 \varphi^*}{\partial \beta^2} \frac{\partial^2 \varphi^*}{\partial \alpha^2} - \frac{\partial^2 \varphi^*}{\partial \beta \partial \alpha} \frac{\partial^2 \varphi^*}{\partial \alpha \partial \beta} \right]_{\alpha=\hat{\alpha}_u, \beta=\hat{\beta}_u}^{-1}. \quad (45)$$

Finally in the above calculation processes, setting  $u(\alpha, \beta)$  as  $\alpha$  and  $\beta$  respectively, the estimates on the basis of SEL function are given by

$$\hat{\alpha}_S = \sqrt{\frac{|\Sigma_\alpha^*|}{|\Sigma|}} \exp\{n[\varphi_\alpha^*(\hat{\alpha}_\alpha, \hat{\beta}_\alpha) - \varphi(\hat{\alpha}_1, \hat{\beta}_1)]\} \quad \text{and} \quad (46)$$

$$\hat{\beta}_S = \sqrt{\frac{|\Sigma_\beta^*|}{|\Sigma|}} \exp\{n[\varphi_\beta^*(\hat{\alpha}_\beta, \hat{\beta}_\beta) - \varphi(\hat{\alpha}_1, \hat{\beta}_1)]\}. \quad (47)$$

Further, we are able to calculate the Bayesian estimates based on BSEL function using the equation  $\delta_{BS} = \sigma \delta_0 + (1 - \sigma) \delta_S$  with different  $\sigma, 0 \leq \sigma \leq 1$ .

Similarly, by treating  $u_1(\alpha, \beta)$  as  $e^{-h\alpha}$  and  $u_2(\alpha, \beta)$  as  $e^{-h\beta}$ , the estimates for unknown parameters under LINEX loss function are given by

$$\hat{\alpha}_L = -\frac{1}{h} \ln \left\{ \sqrt{\frac{|\Sigma_{u_1}^*|}{|\Sigma|}} \exp \{n[\varphi_{u_1}^*(\hat{\alpha}_{u1}, \hat{\beta}_{u1}) - \varphi(\hat{\alpha}_1, \hat{\beta}_1)]\} \right\} \quad \text{and} \quad (48)$$

$$\hat{\beta}_L = -\frac{1}{h} \ln \left\{ \sqrt{\frac{|\Sigma_{u_2}^*|}{|\Sigma|}} \exp \{n[\varphi_{u_2}^*(\hat{\alpha}_{u2}, \hat{\beta}_{u2}) - \varphi(\hat{\alpha}_1, \hat{\beta}_1)]\} \right\}. \quad (49)$$

### 3.3. MH Algorithm

We derive Bayesian estimates for  $\alpha$  and  $\beta$  by the MH algorithm (see Reference [20]). First, we suppose that the bivariate normal distribution is the proposal distribution for the parameter  $\theta = (\alpha, \beta)$ , then the MH algorithm can generate samples from the bivariate normal distribution and finally get convergent samples from the posterior distribution. Using the samples, we first compute the Bayesian estimates under different loss functions, thereafter, establish HPD intervals. The MH algorithm can be summarized as:

- (1) Begin with an initial value  $\theta_0 = (\alpha_0, \beta_0)$ , set  $n = 1$ .
- (2) Generate a proposal  $\theta' = (\alpha', \beta')$  from the bivariate normal distribution  $N_2(\theta_{n-1}, \Sigma_1)$  where  $\theta_{n-1} = (\alpha_{n-1}, \beta_{n-1})$ , and  $\Sigma_1$  denotes the variance-covariance matrix which tends to be considered as the inverse for Fisher information matrix.
- (3) Calculate the acceptance probability  $q = \min \left\{ \frac{\pi(\theta'|X)}{\pi(\theta_{n-1}|X)}, 1 \right\}$ , and  $\pi(\cdot|X)$  is corresponding joint posterior distribution.
- (4) Generate  $\mu$  from Uniform(0, 1).
- (5) If  $\mu \leq q$ , let  $\theta_n = \theta'$ ; else, let  $\theta_n = \theta_{n-1}$ .
- (6) Set  $n = n + 1$ .
- (7) Repeat steps (2–6)  $D$  times to get required size of sample.

Removing the first  $D_0$  number of iterative values, the Bayesian estimates under SEL function are derived as

$$\tilde{\alpha}_S = \frac{1}{D - D_0} \sum_{n=D_0+1}^D \alpha_n \quad \text{and} \quad (50)$$

$$\tilde{\beta}_S = \frac{1}{D - D_0} \sum_{n=D_0+1}^D \beta_n. \quad (51)$$

Proceeding similarly, the desired estimates under BSEL function can be obtained easily. Further the Bayesian estimates under LINEX can be computed as

$$\tilde{\alpha}_L = -\frac{1}{h} \ln \left\{ \frac{1}{D - D_0} \sum_{n=D_0+1}^D e^{-h\alpha_n} \right\} \quad \text{and} \quad (52)$$

$$\tilde{\beta}_L = -\frac{1}{h} \ln \left\{ \frac{1}{D - D_0} \sum_{n=D_0+1}^D e^{-h\beta_n} \right\}. \quad (53)$$

Now, we can establish the  $100(1 - \gamma)\%$  HPD interval (see Reference [25]) for the unknown parameter  $\alpha$ . Sort the remaining  $D - D_0$  values in ascending order to be  $\alpha_{(1)}, \alpha_{(2)}, \dots, \alpha_{(D-D_0)}$ . The  $100(1 - \gamma)\%$  HPD interval of  $\alpha$  is given as

$$(\alpha_{(w^*)}, \alpha_{(w^* + [(1 - \gamma) \times (D - D_0)])}), \quad (54)$$

where  $w^*$  is selected when the following equation is satisfied:

$$\alpha_{(w^* + [(1-\gamma) \times (D - D_0)])} - \alpha_{(w^*)} = \min_{1 \leq w \leq (D - D_0) - [(1-\gamma) \times (D - D_0)]} (\alpha_{(w + [(1-\gamma) \times (D - D_0)])} - \alpha_{(w)}), \quad (55)$$

and  $[(1 - \gamma) \times (D - D_0)]$  denotes the integer part of  $(1 - \gamma) \times (D - D_0)$ . Likewise, the HPD interval of  $\beta$  can be obtained.

#### 4. Bayesian Prediction

Now we obtain the prediction estimates for the future sample on the basis of available sample and obtain predictive intervals. Bayesian prediction for future sample is a fundamental subject in many fields such as medical, agricultural and engineering experiments. Interested readers may refer to Reference [11].

Suppose that the existing  $X = (X_{r+1}, \dots, X_m)$  is a group of general progressive censored data observed from a population with Gompertz distribution. Let  $Y_1 \leq Y_2 \leq \dots \leq Y_W$  denote the ordered failures time for a future sample with size  $W$ , which is also obtained from Gompertz distribution. We aim to obtain their predictive estimation (two-sample prediction). Suppose that  $Y_v$  ( $1 \leq v \leq W$ ) represents the  $v$ -th failure time of the future sample. Then, for given  $\alpha$  and  $\beta$ , we can obtain the density function of  $Y_v$  as

$$\begin{aligned} g(y_v | \alpha, \beta) &= v \binom{W}{v} [1 - F(y_v | \alpha, \beta)]^{W-v} [F(y_v | \alpha, \beta)]^{v-1} f(y_v | \alpha, \beta) \\ &= v \binom{W}{v} \sum_{j=0}^{v-1} \binom{v-1}{j} (-1)^j [1 - F(y_v | \alpha, \beta)]^{W-v+j} f(y_v | \alpha, \beta). \end{aligned} \quad (56)$$

Consequently, the posterior predictive density function for  $Y_v$  is derived as

$$g^*(y_v | X) = \int_0^{+\infty} \int_0^{+\infty} g(y_v | \alpha, \beta) \pi(\alpha, \beta | X) d\alpha d\beta. \quad (57)$$

It is infeasible to compute (57) analytically. By using the MH algorithm, we can obtain its approximate solution as follows:

$$g^*(y_v | X) = \frac{1}{D - D_0} \sum_{i=D_0+1}^D g(y_v | \alpha_i, \beta_i). \quad (58)$$

Further, the survival function is computed as

$$S(y_v | \alpha, \beta) = \int_{y_v}^{+\infty} g(z | \alpha, \beta) dz. \quad (59)$$

The posterior predictive survival function for  $Y_v$  can be derived by

$$\begin{aligned} S'_1(y_v | X) &= \int_0^{+\infty} \int_0^{+\infty} S(y_v | \alpha, \beta) \pi(\alpha, \beta | X) d\alpha d\beta \\ &\approx \frac{1}{D - D_0} \sum_{i=D_0+1}^D \int_{y_v}^{+\infty} g(z | \alpha_i, \beta_i) dz \\ &= \frac{v}{D - D_0} \binom{W}{v} \sum_{i=D_0+1}^D \sum_{j=0}^{v-1} \binom{v-1}{j} (-1)^j \frac{[1 - F(y_v | \alpha_i, \beta_i)]^{W+j-v+1}}{W+j-v+1}. \end{aligned} \quad (60)$$

Then, we construct the  $100(1 - \gamma)\%$  Bayesian predictive interval  $(L_0, U_0)$  of  $Y_v$  by finding the solution of the equations

$$S'_1(L_0 | X) = 1 - \frac{\gamma}{2} \quad \text{and} \quad S'_1(U_0 | X) = \frac{\gamma}{2}. \quad (61)$$

Further, it is convenient to derive the predictive estimate of the future  $v$ -th ordered lifetime, which is given by

$$\begin{aligned}\hat{y}_v &= E(y_v|X) = \int_0^{+\infty} y_v g^*(y_v|X) dy_v \\ &= \int_0^{+\infty} \int_0^{+\infty} H(\alpha, \beta) \pi(\alpha, \beta|X) d\alpha d\beta,\end{aligned}\quad (62)$$

where  $H(\alpha, \beta) = \int_0^{+\infty} y_v g(y_v|\alpha, \beta) dy_v$  is obtained as

$$H(\alpha, \beta) = \frac{v\alpha}{\beta} \binom{W}{v} \sum_{j=0}^{v-1} \binom{v-1}{j} (-1)^j \int_0^{+\infty} e^{-\alpha(W+1-v+j)t} \ln(1+t) dt. \quad (63)$$

By using the MH algorithm described in the previous section, the prediction estimate of  $Y_k$  is derived as

$$\hat{y}_v = \frac{1}{D - D_0} \sum_{i=D_0+1}^D H(\alpha_i, \beta_i). \quad (64)$$

## 5. Simulation and Data Analysis

For evaluating the quality of the approaches, a numeric simulation study is carried out. In addition, we also analyze two real data sets for further illustration.

### 5.1. Simulation Study

For the sake of simulation, first we generate general progressive censored sample with the algorithm discussed by Reference [22]. The procedures are as below:

- (1) Generate  $B_m$  from  $\text{Beta}(n - r, r + 1)$ .
- (2) Generate independent  $U_{r+k}$  from  $\text{Uniform}(0, 1)$ ,  $k = 1, 2, \dots, m - r - 1$ .
- (3) Set  $B_{r+k} = U_{r+k}^{\frac{1}{\xi_{r+k}}}$ , where  $\xi_{r+k} = k + \sum_{i=m-k+1}^m R_i$ ,  $k = 1, 2, \dots, m - r - 1$ .
- (4) Set  $Z_{r+k} = 1 - B_{m-k+1} B_{m-k+2} \dots B_m$ ,  $k = 1, 2, \dots, m - r$ .
- (5) Set  $X_k = F^{-1}(Z_k)$ ,  $k = r + 1, \dots, m$ , and  $F(x)$  represents the CDF of Gompertz distribution.

Then we get the desired general progressive censored data  $X_i, i = r + 1, \dots, m$  drawn from Gompertz distribution. In our experiment, the true values for  $(\alpha, \beta)$  are selected to be  $(0.3, 1.2)$ . The MLEs for the two parameters are calculated by means of EM algorithm. In the aspect of Bayesian estimation and prediction,  $(0.2, 7.8, 0.1, 3.7)$  are chosen to be the values of hyperparameters  $(a, b, c, d)$  respectively. Moreover, Bayesian estimates are obtained by TK and MH methods under different loss functions. Comparison between the results is made according to mean-square error (MSE).

For convenience, we use simplified notations to represent different censoring schemes (CS) with  $r$ , such as  $(2, 0^{*3}, 5)$  for the case where  $r = 2$ ,  $m = 6$ ,  $n = 10$  and the censoring scheme is  $(0, 0, 0, 5)$ . Therefore, our schemes in simulation study can be expressed by the following notations:  $H_1 = (0^{*2}, 3, 0^{*3}, 4, 2, 0^{*4})$ ,  $H_2 = (2, 0^{*3}, 2, 0^{*5}, 5, 0)$ ,  $H_3 = (2, 2, 0^{*5}, 8, 0^{*7}, 3)$ ,  $H_4 = (5, 0^{*2}, 1, 2, 0^{*3}, 7, 0^{*7})$ ,  $H_5 = (3, 5, 0^{*4}, 2, 0^{*8}, 10, 0^{*5})$ ,  $H_6 = (5, 0^{*4}, 5, 0^{*3}, 4, 0^{*4}, 6, 0^{*6})$ ,  $H_7 = (0^{*4}, 6, 0^{*3}, 7, 0^{*9}, 2, 0^{*7})$ ,  $H_8 = (3, 0^{*2}, 2, 0^{*4}, 10, 0^{*4}, 5, 0^{*12})$ ,  $H_9 = (2, 0^{*5}, 10, 0^{*5}, 6, 0^{*7}, 2, 0^{*10})$ ,  $H_{10} = (5, 0, 3, 0^{*8}, 8, 0^{*6}, 4, 0^{*12})$ .

Table 1 reports all the average estimates and corresponding MSEs for the parameters. In this table, for a given censoring scheme, the average estimates are placed on the first and third rows respectively, and the second and fourth rows refer to the corresponding MSEs. From tabulated estimates, in general, the MH estimates are observed to have smaller MSEs compared with the estimates using the TK method. Furthermore, we find that the performance of Bayes estimates for the parameters under LINEX is better than those based on the SEL and BSEL functions in MSEs. However, Bayesian estimates under the SEL function are closer to the actual values. For MLEs, it can be seen that larger  $m - r$  and  $n$  bring about more outstanding estimates, where  $m - r$  and  $n$  are the corresponding sizes of observed and complete sample. On the whole, Bayesian estimates have an advantage over the corresponding MLEs.

Furthermore, different intervals have also been constructed, including ACIs on the basis of Fisher information matrix, parametric bootstrap intervals, and HPD intervals based on the sample generated from the MH algorithm. Table 2 presents their average length (AL) and coverage probabilities (CPs). The tabulated values indicate that the AL of the HPD intervals is the shortest among those obtained from other interval estimates. Besides, we also find that the ACIs have better performance according to CPs. In general, bootstrap-t and bootstrap-p intervals behave similarly, and their CPs tend to be below the 95% confidence level. Table 3 lists the results of point prediction and 95% interval prediction. We give the prediction results of  $y_3$ ,  $y_7$  and  $y_{10}$  in a future sample with size 10. Furthermore, we discover that the interval length becomes wider as  $v$  increases.

## 5.2. Data Analysis

**Dataset 1:** First we analyze a real dataset about the breaking stress of carbon fibers (in Gba) ( $n = 66$ ) (see Reference [26]). It is listed as follows:

3.70, 2.74, 2.73, 2.50, 3.60, 3.11, 3.27, 2.87, 1.47, 3.11, 3.56, 4.42, 2.41, 3.19, 3.22, 1.69, 3.28, 3.09, 1.87, 3.15, 4.90, 1.57, 2.67, 2.93, 3.22, 3.39, 2.81, 4.20, 3.33, 2.55, 3.31, 3.31, 2.85, 1.25, 4.38, 1.84, 0.39, 3.68, 2.48, 0.85, 1.61, 2.79, 4.70, 2.03, 1.89, 2.88, 2.82, 2.05, 3.65, 3.75, 2.43, 2.95, 2.97, 3.39, 2.96, 2.35, 2.55, 2.59, 2.03, 1.61, 2.12, 3.15, 1.08, 2.56, 1.80, 2.53.

In order to analyze these data, we calculate the MLEs for the two parameters, and then for the Gompertz distribution we conduct a goodness-of-fit test with some practical guidelines like the Kolmogorov-Smirnov (K-S) statistics, and other criteria, for example, the Akaike Information Criterion (AIC), as well as the Bayesian Information Criterion (BIC). For comparison, some other life distributions have also been tested for goodness-of-fit, including Generalized Exponential (GE), Inverse Weibull and Exponential distributions. Their PDFs have the following forms, respectively:

(1) The PDF of GE distribution:

$$f_{GE}(x|\alpha, \beta) = \alpha\beta(1 - e^{-\beta x})^{\alpha-1}e^{-\beta x}, \quad 0 < x < +\infty, 0 < \alpha, \beta < +\infty;$$

(2) The PDF of Inverse Weibull distribution:

$$f_{IW}(x|\alpha, \beta) = \alpha\beta x^{-\alpha-1}e^{-\beta x^{-\alpha}}, \quad 0 < x < +\infty, 0 < \alpha, \beta < +\infty;$$

(3) The PDF of Exponential distribution:

$$f_E(x|\beta) = \beta e^{-\beta x}, \quad 0 < x < +\infty, 0 < \beta < +\infty.$$

**Table 1.** Estimates and MSEs under general progressive censoring schemes.

$n(r, m - r)$	CS	TK						MH					
		BSEL			LINEX			BSEL			LINEX		
		MLE	SEL	$\sigma = 0.3$	$\sigma = 0.7$	$h = -2$	$h = 2$	SEL	$\sigma = 0.3$	$\sigma = 0.7$	$h = -2$	$h = 2$	
20(0,11)	$H_1$	0.3658 (0.3800)	0.3108 (0.004819)	0.3273 (0.04809)	0.3493 (0.1982)	0.3589 (0.03838)	0.2613 (0.009074)	0.3130 (0.0078843)	0.3288 (0.003863)	0.3500 (0.0007096)	0.3294 (0.008166)	0.2894 (0.005391)	
		1.581 (0.5589)	1.180 (0.03178)	1.461 (0.1174)	1.443 (0.3282)	1.329 (0.07918)	1.048 (0.02627)	1.180 (0.03099)	1.300 (0.01518)	1.461 (0.002789)	1.353 (0.002789)	1.104 (0.002788)	
20(2,11)	$H_2$	0.3642 (0.2977)	0.3188 (0.004574)	0.3324 (0.03937)	0.3505 (0.1566)	0.4262 (0.0388)	0.2172 (0.007151)	0.3236 (0.005837)	0.3358 (0.002860)	0.3520 (0.0005253)	0.3328 (0.006824)	0.2864 (0.005360)	
		1.602 (0.6573)	1.170 (0.02902)	1.300 (0.1272)	1.472 (0.3785)	1.412 (0.07533)	0.9907 (0.02982)	1.157 (0.03075)	1.291 (0.01507)	1.469 (0.002768)	1.372 (0.002768)	1.067 (0.002768)	
30(2,15)	$H_3$	0.3541 (0.2357)	0.3231 (0.005198)	0.3324 (0.03374)	0.3448 (0.1259)	0.3455 (0.03671)	0.2880 (0.005695)	0.3212 (0.007422)	0.3311 (0.003637)	0.3442 (0.0006680)	0.3414 (0.007797)	0.3013 (0.004011)	
		1.479 (0.3951)	1.193 (0.03355)	1.279 (0.09828)	1.393 (0.2429)	1.330 (0.05929)	1.080 (0.03756)	1.186 (0.03513)	1.274 (0.01721)	1.391 (0.003161)	1.311 (0.003161)	1.104 (0.003161)	
30(5,15)	$H_4$	0.3375 (0.1270)	0.3189 (0.006982)	0.3245 (0.02480)	0.3319 (0.07280)	0.3846 (0.04221)	0.2314 (0.005941)	0.3124 (0.007436)	0.3199 (0.003644)	0.3300 (0.0006692)	0.3315 (0.009107)	0.2863 (0.006218)	
		1.408 (0.2745)	1.225 (0.04216)	1.280 (0.08951)	1.353 (0.1825)	1.387 (0.05769)	1.089 (0.06132)	1.216 (0.03791)	1.274 (0.01858)	1.351 (0.003412)	1.316 (0.003412)	1.175 (0.003412)	
40(3,20)	$H_5$	0.3260 (0.08369)	0.3225 (0.007040)	0.3236 (0.01979)	0.3250 (0.05045)	0.2999 (0.03917)	0.3370 (0.004908)	0.3210 (0.007466)	0.3225 (0.003658)	0.3263 (0.0006720)	0.3263 (0.01124)	0.3030 (0.008610)	
		1.382 (0.2233)	1.208 (0.03806)	1.260 (0.07655)	1.330 (0.1507)	1.265 (0.05024)	1.161 (0.05376)	1.206 (0.03845)	1.259 (0.01884)	1.329 (0.003460)	1.319 (0.003460)	1.147 (0.003460)	
40(5,20)	$H_6$	0.3270 (0.07880)	0.3239 (0.007134)	0.3248 (0.01915)	0.3261 (0.04781)	0.3360 (0.04173)	0.2854 (0.004668)	0.3140 (0.009779)	0.3180 (0.004792)	0.3231 (0.0008801)	0.3268 (0.007355)	0.2942 (0.007396)	
		1.350 (0.1814)	1.211 (0.03611)	1.253 (0.06730)	1.308 (0.1254)	1.305 (0.04621)	1.135 (0.06150)	1.233 (0.04281)	1.268 (0.02098)	1.315 (0.003853)	1.292 (0.004146)	1.169 (0.004146)	
40(0,25)	$H_7$	0.3153 (0.05276)	0.3096 (0.007417)	0.3113 (0.01595)	0.3136 (0.03409)	0.4376 (0.05085)	0.1756 (0.004505)	0.3026 (0.007637)	0.3064 (0.003742)	0.3115 (0.0006873)	0.3075 (0.007945)	0.3023 (0.006896)	
		1.341 (0.1537)	1.238 (0.03891)	1.269 (0.06481)	1.310 (0.1107)	1.415 (0.03442)	1.045 (0.06736)	1.240 (0.04114)	1.271 (0.02016)	1.311 (0.003702)	1.326 (0.003702)	1.153 (0.003702)	
40(3,25)	$H_8$	0.3202 (0.03832)	0.3189 (0.007207)	0.3192 (0.01349)	0.3198 (0.02594)	0.2732 (0.04694)	0.3418 (0.003930)	0.3010 (0.008018)	0.3067 (0.003929)	0.3144 (0.0007216)	0.3339 (0.01114)	0.3050 (0.006569)	
		1.299 (0.1220)	1.216 (0.03458)	1.241 (0.05478)	1.274 (0.08973)	1.244 (0.03643)	1.216 (0.06051)	1.215 (0.03461)	1.240 (0.01696)	1.274 (0.003115)	1.306 (0.003115)	1.159 (0.003115)	
50(2,30)	$H_9$	0.3130 (0.04112)	0.3123 (0.008172)	0.3125 (0.01486)	0.3128 (0.02804)	0.3868 (0.04575)	0.2266 (0.003879)	0.3166 (0.01427)	0.3155 (0.006994)	0.3141 (0.001285)	0.3200 (0.01180)	0.3026 (0.01074)	
		1.310 (0.1154)	1.234 (0.03679)	1.257 (0.05546)	1.287 (0.08691)	1.364 (0.02824)	1.104 (0.07838)	1.223 (0.05966)	1.249 (0.02923)	1.284 (0.005369)	1.299 (0.005369)	1.182 (0.005369)	
50(5,30)	$H_{10}$	0.3079 (0.03577)	0.3129 (0.007717)	0.3114 (0.01351)	0.3094 (0.02473)	0.3298 (0.04378)	0.3012 (0.002980)	0.3120 (0.006585)	0.3108 (0.003226)	0.3091 (0.0005926)	0.3127 (0.008346)	0.2990 (0.007515)	
		1.313 (0.1103)	1.238 (0.03658)	1.261 (0.05427)	1.291 (0.08375)	1.280 (0.02626)	1.173 (0.06414)	1.219 (0.03330)	1.247 (0.01631)	1.285 (0.002997)	1.311 (0.002997)	1.194 (0.002997)	

**Table 2.** Interval estimates with confidence level of 95% for  $\alpha$  and  $\beta$ .

$n(r, m - r)$	CS	Asymptotic		Bootstrap-p		Bootstrap-t		HPD	
		ALs	CPs	ALs	CPs	ALs	CPs	ALs	CPs
20(0,11)	$H_1$	2.060	0.77	1.570	0.85	1.551	0.84	0.4837	0.83
		2.458	0.94	2.839	0.79	2.911	0.80	1.388	0.78
20(2,11)	$H_2$	3.153	0.73	2.202	0.88	1.641	0.84	0.4582	0.87
		2.907	0.95	3.189	0.82	2.903	0.85	1.186	0.81
30(2,15)	$H_3$	3.391	0.78	1.691	0.83	1.979	0.82	0.6035	0.87
		2.346	0.94	2.552	0.87	2.664	0.84	1.139	0.84
30(5,15)	$H_4$	1.006	0.83	0.9996	0.84	0.9396	0.80	0.4921	0.84
		1.753	0.95	1.944	0.81	1.885	0.86	0.8988	0.86
40(3,20)	$H_5$	1.001	0.83	0.9348	0.84	0.8464	0.85	0.6066	0.85
		1.755	0.95	1.874	0.81	1.740	0.83	0.5879	0.84
40(5,20)	$H_6$	0.8690	0.83	0.8693	0.84	0.9051	0.88	0.4502	0.86
		1.596	0.95	1.735	0.86	1.770	0.84	0.9282	0.83
40(0,25)	$H_7$	0.9111	0.85	0.8762	0.87	0.9412	0.83	0.2881	0.74
		1.450	0.93	1.520	0.82	1.488	0.85	0.8589	0.91
40(3,25)	$H_8$	0.7631	0.84	0.7806	0.79	0.8429	0.85	0.3531	0.93
		1.354	0.93	1.431	0.85	1.465	0.80	0.8237	0.92
50(2,30)	$H_9$	0.7248	0.86	0.7310	0.88	0.5984	0.82	0.4184	0.99
		1.257	0.95	1.315	0.87	1.283	0.84	0.6921	0.96
50(5,30)	$H_{10}$	0.6995	0.85	0.7021	0.85	0.7254	0.84	0.4984	0.83
		1.227	0.94	1.285	0.83	1.308	0.88	0.8684	0.87

**Table 3.** Point prediction and 95% prediction interval with  $W = 10$ .

$n(r, m - r)$	CS	$v$	Point Prediction	Interval Prediction
20(0,11)	$H_1$	3	0.6797	$(7.802 \times 10^{-5}, 1.991)$
		7	1.437	$(0.6615, 3.181)$
		10	2.407	$(1.954, 4.649)$
20(2,11)	$H_2$	3	0.7784	$(6.630 \times 10^{-5}, 2.222)$
		7	1.665	$(1.007, 3.402)$
		10	2.717	$(2.345, 4.952)$
30(2,15)	$H_3$	3	0.6801	$(4.705 \times 10^{-5}, 2.063)$
		7	1.474	$(0.8233, 3.029)$
		10	2.349	$(2.157, 4.415)$
30(5,15)	$H_4$	3	0.6176	$(6.116 \times 10^{-5}, 1.881)$
		7	1.325	$(0.9165, 2.787)$
		10	2.155	$(2.011, 3.876)$
40(3,20)	$H_5$	3	0.6123	$(4.757 \times 10^{-5}, 1.925)$
		7	1.369	$(0.8673, 2.851)$
		10	2.094	$(2.035, 3.937)$
40(5,20)	$H_6$	3	0.7092	$(4.432 \times 10^{-5}, 2.074)$
		7	1.445	$(1.059, 2.968)$
		10	2.263	$(2.221, 4.123)$
40(0,25)	$H_7$	3	0.6147	$(5.143 \times 10^{-5}, 1.905)$
		7	1.318	$(0.8493, 2.812)$
		10	2.089	$(2.032, 3.951)$
40(3,25)	$H_8$	3	0.6341	$(5.932 \times 10^{-5}, 1.986)$
		7	1.320	$(1.021, 2.838)$
		10	2.096	$(2.106, 3.902)$
50(2,30)	$H_9$	3	0.6090	$(7.831 \times 10^{-5}, 1.914)$
		7	1.325	$(1.033, 2.783)$
		10	2.025	$(2.092, 3.867)$
50(5,30)	$H_{10}$	3	0.5322	$(4.538 \times 10^{-5}, 1.859)$
		7	1.210	$(1.021, 2.651)$
		10	1.918	$(2.025, 3.665)$

The results of the tests are presented in Table 4 together with the MLEs. Note that the distribution is more suitable to fit the data when K-S, AIC and BIC values are smaller. Comparing the values, we can conclude that the Gompertz distribution is more appropriate.

To illustrate the proposed methods, three groups of general progressive censored data have been randomly drawn from the parent sample as follows:

**Scheme 1:**  $(-, -, 1.25, 1.47, 1.57, 1.61, 1.61, 1.69, 1.80, 1.87, 2.03, 2.05, 2.35, 2.41, 2.43, 2.48, 2.50, 2.53, 2.55, 2.56, 2.67, 2.73, 2.97, 3.11, 3.15, 3.22, 4.42)$ ,  $r = 3$ ,  $m - r = 25$ ,  $R_6 = 11$ ,  $R_{15} = 8$ ,  $R_{19} = 14$ ,  $R_{26} = 5$ ,  $R_i = 0$ , others;

**Scheme 2:**  $(-, -, -, 1.57, 1.61, 1.61, 1.69, 1.80, 1.84, 1.87, 2.03, 2.03, 2.05, 2.12, 2.43, 2.48, 2.55, 2.59, 2.67, 2.73, 2.82, 2.87, 2.88, 2.96, 3.09, 3.11, 3.11, 3.15, 3.19, 3.60, 3.75, 4.42, 4.70)$ ,  $r = 5$ ,  $m - r = 30$ ,  $R_6 = 4$ ,  $R_{12} = 13$ ,  $R_{16} = 10$ ,  $R_{30} = 4$ ,  $R_i = 0$ , others;

**Scheme 3:**  $(-, -1.08, 1.25, 1.47, 1.57, 1.61, 1.61, 1.69, 1.80, 1.84, 1.89, 2.03, 2.03, 2.05, 2.12, 2.35, 2.41, 2.48, 2.50, 2.53, 2.55, 2.79, 2.82, 2.93, 2.95, 3.19, 3.22, 3.27, 3.31, 3.33, 3.39, 3.60, 3.68, 3.75, 4.20, 4.90)$ ,  $r = 2$ ,  $m - r = 35$ ,  $R_4 = 7$ ,  $R_{17} = 9$ ,  $R_{22} = 13$ ,  $R_i = 0$ , others.

**Table 4.** The MLE and goodness-of-fit tests results in Dataset 1.

Distribution	$\hat{\alpha}_{\text{ML}}$	$\hat{\beta}_{\text{ML}}$	K-S	AIC	BIC
Gompertz	0.0348201	1.07068	0.11122	180.177	184.556
GE	9.19911	1.00755	0.15472	194.745	199.124
Inverse weibull	1.64805	3.22624	0.23042	246.390	250.769
Exponential		0.362379	0.28615	267.989	270.178

With the EM algorithm, we calculate the MLEs, and the corresponding Bayesian estimates are also derived by TK and MH methods. For the sake of no prior information, all the hyperparameters are set close to zero values. We list the average MLEs and Bayesian estimates in Tables 5 and 6. In Table 7, the 90% interval estimates are tabulated, which are ACIs, parametric bootstrap and HPD intervals. Finally in Table 8, point prediction and 95% interval prediction of  $y_1$  and  $y_6$  in a future sample with size 6 are presented.

**Table 5.** Estimates for  $\alpha$  and  $\beta$  by EM and TK method in Dataset 1.

CS	MLE	SEL	BSEL				LINEX			
			$\sigma = 0.2$	$\sigma = 0.4$	$\sigma = 0.6$	$\sigma = 0.8$	$h = -5$	$h = -3$	$h = 5$	$h = 7$
1	0.01946	0.02609	0.02476	0.02344	0.02211	0.02079	0.02346	0.02923	0.006137	0.008611
	1.227	1.199	1.205	1.210	1.216	1.221	1.290	1.142	1.734	1.671
2	0.03102	0.03786	0.03649	0.03512	0.03376	0.03239	0.03635	0.04326	0.01559	0.01855
	1.091	1.072	1.076	1.079	1.083	1.087	1.126	1.035	1.397	1.358
3	0.02936	0.03555	0.03431	0.03307	0.03183	0.03060	0.03419	0.04052	0.01519	0.01790
	1.091	1.074	1.077	1.081	1.084	1.088	1.123	1.039	1.377	1.340

**Table 6.** Estimates for  $\alpha$  and  $\beta$  by MH algorithm in Dataset 1.

CS	SEL	BSEL				LINEX			
		$\sigma = 0.2$	$\sigma = 0.4$	$\sigma = 0.6$	$\sigma = 0.8$	$h = -5$	$h = -3$	$h = 5$	$h = 7$
1	0.03833	0.03456	0.03078	0.02701	0.02323	0.02529	0.02813	0.02342	0.02547
	1.061	1.095	1.128	1.161	1.194	1.264	1.228	1.145	1.090
2	0.04482	0.04206	0.03930	0.03654	0.03378	0.03903	0.03765	0.03707	0.03750
	1.013	1.029	1.044	1.060	1.075	1.115	1.095	1.020	0.9947
3	0.05173	0.04725	0.04278	0.03831	0.03383	0.03625	0.03591	0.03631	0.03775
	0.9878	1.008	1.029	1.050	1.071	1.114	1.101	1.022	0.9948

**Table 7.** Interval estimates with confidence level of 90% for  $\alpha$  and  $\beta$  in Dataset 1.

CS	Asymptotic	Bootstrap-p	Bootstrap-t	HPD
1	(0.001019, 0.03790) (0.9007, 1.554)	(0.004187, 0.04863) (0.9560, 1.740)	(0.004368, 0.04738) (0.9548, 1.743)	(0.01946, 0.06289) (0.7067, 1.227)
	(0.004108, 0.05794) (0.8435, 1.338)	(0.009230, 0.06913) (0.8692, 1.464)	(0.01021, 0.07103) (0.8682, 1.439)	(0.02092, 0.09859) (0.7160, 1.197)
2	(0.004407, 0.05431) (0.8534, 1.329)	(0.009369, 0.06255) (0.8871, 1.436)	(0.009445, 0.06360) (0.8900, 1.434)	(0.01217, 0.04114) (0.9851, 1.321)

**Table 8.** Point prediction and 95% interval prediction with  $W = 6$  in Dataset 1.

CS	$v$	Point Prediction	Interval Prediction
1	1	1.309	$(7.982 \times 10^{-5}, 3.763)$
	6	3.985	$(3.898, 5.566)$
2	1	1.414	$(4.707 \times 10^{-5}, 3.619)$
	6	4.078	$(4.004, 5.696)$
3	1	1.457	$(4.672 \times 10^{-5}, 3.666)$
	6	3.833	$(3.817, 6.542)$

**Dataset 2:** Reference [27] presented a dataset on the tumor-free days of 30 rats which were fed with unsaturated diet which is listed below as

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.

In order to analyze these data, Reference [28] assumed that the number of tumor-free days obeys the Gompertz distribution. To illustrate the methods discussed, here we also suppose that the distribution for these data is Gompertz with  $(\alpha, \beta)$ . Let  $m - r = 20$ , we set up two censoring schemes, respectively  $s'_1 = (0^{*4}, 5, 0^{*7}, 2, 0^{*7})$ ,  $r = 3$  and  $s'_2 = (0^{*2}, 4, 0^{*6}, 5, 0^{*10})$ ,  $r = 1$ . Then we have obtained the sample under  $s'_1$ :

$$63, 66, 66, 66, 68, 70, 70, 77, 77, 84, 91, 91, 94, 98, 101, 105, 108, 112, 115, 178,$$

and the sample under  $s'_2$ :

$$63, 63, 63, 66, 66, 66, 68, 70, 70, 77, 77, 84, 91, 91, 94, 101, 109, 112, 115, 178.$$

In Tables 9 and 10, we calculate the average MLEs, and average Bayesian estimates are derived by the TK method and the MH algorithm, respectively. The interval estimates are presented in Table 11, including ACIs, parametric bootstrap and HPD intervals. Finally Table 12 presents the results of the point prediction and the 95% interval prediction of  $y_1$  and  $y_5$  with  $W = 5$ .

**Table 9.** Estimates for  $\alpha$  and  $\beta$  by EM and TK methods in Dataset 2.

CS	MLE	SEL	BSEL				LINEX			
			$\sigma = 0.2$	$\sigma = 0.4$	$\sigma = 0.6$	$\sigma = 0.8$	$h = -3$	$h = -1$	$h = 5$	$h = 7$
$s'_1$	0.08360	0.1125	0.1067	0.1009	0.09516	0.08938	0.07764	0.1013	0.05868	0.06071
	0.02461	.02383	0.02398	0.02414	0.02430	0.02446	0.02546	0.02360	0.02696	0.02680
$s'_2$	0.07455	0.09943	0.09446	0.08948	0.08450	0.07953	0.06870	0.08973	0.05189	0.05369
	0.02526	0.02449	0.02464	0.02480	0.02495	0.02510	0.02612	0.02426	0.02761	0.02745

**Table 10.** Estimates for  $\alpha$  and  $\beta$  by MH algorithm in Dataset 2.

CS	SEL	BSEL				LINEX			
		$\sigma = 0.2$	$\sigma = 0.4$	$\sigma = 0.6$	$\sigma = 0.8$	$h = -3$	$h = -1$	$h = 5$	$h = 7$
$s'_1$	0.1171	0.1104	0.1037	0.09701	0.09031	0.09774	0.1044	0.09331	0.09782
	0.02331	0.02357	0.02383	0.02409	0.02435	0.02446	0.02401	0.02426	0.02381
$s'_2$	0.1045	0.09850	0.09251	0.08653	0.08054	0.09351	0.09069	0.08440	0.08858
	0.02405	0.02429	0.02453	0.02478	0.02502	0.02428	0.02447	0.02477	0.02445

**Table 11.** Interval estimates with confidence level of 90% for  $\alpha$  and  $\beta$  in Dataset 2.

CS	Asymptotic	Bootstrap-p	Bootstrap-t	HPD
$s'_1$	(0.001083, 0.1661)	(0.02024, 0.2246)	(0.01836, 0.2146)	(0.007255, 0.1017)
	(0.01715, 0.03207)	(0.01784, 0.03746)	(0.01766, 0.03789)	(0.02008, 0.04409)
$s'_2$	(0.0008985, 0.1482)	(0.01639, 0.1995)	(0.01641, 0.2031)	(0.02321, 0.1732)
	(0.01781, 0.03270)	(0.01810, 0.03904)	(0.01820, 0.03875)	(0.01878, 0.03419)

**Table 12.** Point prediction and 95% interval prediction with  $W = 5$  in Dataset 2.

CS	$v$	Point Prediction	Interval Prediction
$s'_1$	1	41.53805	$(6.196145 \times 10^{-5}, 79.12604)$
	5	136.6431	$(74.30713, 188.8303)$
$s'_2$	1	43.12672	$(5.896611 \times 10^{-5}, 94.66528)$
	5	126.0958	$(86.76296, 169.7341)$

## 6. Conclusions

In summary, we discuss the classical and Bayes inferences for the Gompertz distribution using the general progressive censoring. First, the MLEs are acquired by the Expectation–Maximization algorithm. Then, according to the asymptotic normality of

MLEs and the principle of missing information, we provide the asymptotic confidence intervals. Moreover, we derive parametric percentile bootstrap and bootstrap-t intervals. In Bayesian statistics, three loss functions are considered, which are symmetrical, asymmetrical and balanced, respectively. Since the posterior expectation is intractable to obtain in explicit form, the TK method is employed to calculate approximate Bayesian estimates. Besides, the Metropolis-Hastings algorithm is applied to get the Bayesian estimates and establish HPD intervals. Furthermore, we derive the prediction estimates of future samples. Finally, a numerical simulation is executed to appraise the quality of the approaches, and two real data sets are also analyzed. The results indicate that these approaches have good performance. In addition, the methods in this article can be extended to other distributions.

**Author Contributions:** Investigation, Y.W.; Supervision, W.G. Both authors have read and agreed to the published version of the manuscript.

**Funding:** This research was supported by Project 202110004106 supported by Beijing Training Program of Innovation and Entrepreneurship for Undergraduates.

**Conflicts of Interest:** The authors declare no conflict of interest.

## References

1. Gompertz, B. On the nature of the function expressive of the law of human mortality and on a new mode of determining life contingencies. *Philos. Trans. R. Soc. Lond.* **1825**, *115*, 513–585.
2. Willekens, F. Gompertz in context: The Gompertz and related distributions. In *Forecasting Mortality in Developed Countries: Insights from a Statistical, Demographic and Epidemiological Perspective*; Springer: Berlin/Heidelberg, Germany, 2001; Volume 9, pp. 105–126.
3. Wu, J.-W.; Hung, W.-L.; Tsai, C.-H. Estimation of parameters of the Gompertz distribution using the least squares method. *Appl. Math. Comput.* **2004**, *158*, 133–147. [\[CrossRef\]](#)
4. Chang, S.; Tsai, T. Point and interval estimations for the Gompertz distribution under progressive Type-II censoring. *Metron* **2003**, *61*, 403–418.
5. Mohie El-Din, M.M.; Nagy, M.; Abu-Moussa, M.H. Estimation and prediction for Gompertz distribution under the generalized progressive hybrid censored data. *Ann. Data Sci.* **2019**, *6*, 673–705. [\[CrossRef\]](#)
6. Soliman, A.A.; Abd-Ellah, A.H.; Abou-Elhaggag, N.A. Abd-Elmougod, G.A. Estimation of the parameters of life for Gompertz distribution using progressive first-failure censored data. *Comput. Stat. Data Anal.* **2012**, *56*, 2471–2485. [\[CrossRef\]](#)
7. Bakouch, H.S.; El-Bar, A. A new weighted Gompertz distribution with applications to reliability data. *Appl. Math.* **2017**, *62*, 269–296. [\[CrossRef\]](#)
8. Ghitany, M.; Alqallaf, F.; Balakrishnan, N. On the likelihood estimation of the parameters of Gompertz distribution based on complete and progressively Type-II censored samples. *J. Stat. Comput. Simul.* **2014**, *84*, 1803–1812. [\[CrossRef\]](#)
9. Balakrishnan, N.; Sandhu, R. Best linear unbiased and maximum likelihood estimation for exponential distributions under general progressive Type-II censored samples. *Sankhyā Indian J. Stat. Ser.* **1996**, *58*, 1–9.
10. Fernandez, A.J. On estimating exponential parameters with general Type-II progressive censoring. *J. Stat. Plan. Inference* **2004**, *121*, 135–147. [\[CrossRef\]](#)
11. Peng, X.Y.; Yan, Z.Z. Bayesian estimation and prediction for the Inverse Weibull distribution under general progressive censoring. *Commun. Stat. Theory Methods* **2016**, *45*, 621–635.
12. Soliman, A.A.; Al-Hossain, A.Y.; Al-Harbi, M.M. Predicting observables from Weibull model based on general progressive censored data with asymmetric loss. *Stat. Methodol.* **2011**, *8*, 451–461. [\[CrossRef\]](#)
13. Kim, C.; Han, K. Estimation of the scale parameter of the Rayleigh distribution under general progressive censoring. *J. Korean Stat. Soc.* **2009**, *38*, 239–246. [\[CrossRef\]](#)
14. Soliman, A.A. Estimations for Pareto model using general progressive censored data and asymmetric loss. *Commun. Stat. Theory Methods* **2008**, *37*, 1353–1370. [\[CrossRef\]](#)
15. Wang, B.X. Exact interval estimation for the scale family under general progressive Type-II censoring. *Commun. Stat. Theory Methods* **2012**, *41*, 4444–4452. [\[CrossRef\]](#)
16. Dempster, A.P.; Laird, N.M.; Rubin, D.B. Maximum likelihood from incomplete data via the EM algorithm. *J. R. Stat. Soc. Ser. (Methodol.)* **1977**, *39*, 1–38.
17. Louis, T.A. Finding the observed information matrix when using the EM algorithm. *J. R. Stat. Soc. Ser. (Methodol.)* **1982**, *44*, 226–233.
18. Wang, J.; Wang, X.R. The EM algorithm for the estimation of parameters under the general Type-II progressive censoring data. *J. Anhui Norm. Univ. (Nat. Sci.)* **2014**, *37*, 524–529.
19. Efron, B.; Tibshirani, R. Bootstrap methods for standard errors, confidence intervals, and other measures of statistical accuracy. *Stat. Sci.* **1986**, *1*, 54–75. [\[CrossRef\]](#)

20. Kayal, T.; Tripathi, Y.M.; Singh, D.P.; Rastogi, M.K. Estimation and prediction for Chen distribution with bathtub shape under progressive censoring. *J. Stat. Comput. Simul.* **2017**, *87*, 348–366. [[CrossRef](#)]
21. Kundu, D.; Kannan, N.; Balakrishnan, N. Analysis of progressively censored competing risks data. Advances in Survival Analysis. In *Handbook of Statistics*; Elsevier: New York, NY, USA, 2004; Volume 23, pp. 331–348.
22. Aggarwala, R.; Balakrishnan, N. Some properties of progressive censored order statistics from arbitrary and uniform distributions with applications to inference and simulation. *J. Stat. Plan. Inference* **1998**, *70*, 35–49. [[CrossRef](#)]
23. Tierney, L.; Kadane, J.B. Accurate approximations for posterior moments and marginal densities. *J. Am. Stat. Assoc.* **1986**, *81*, 82–86. [[CrossRef](#)]
24. Jozani, M.J.; March, É.; Parsian, A. Bayesian and robust Bayesian analysis under a general class of balanced loss functions. *Stat. Pap.* **2012**, *53*, 51–60. [[CrossRef](#)]
25. Bai, X.; Shi, Y.; Liu, Y.; Liu, B. Reliability estimation of stress–strength model using finite mixture distributions under progressively interval censoring. *J. Comput. Appl. Math.* **2019**, *348*, 509–524. [[CrossRef](#)]
26. Nichols, M.D.; Padgett, W. A bootstrap control chart for Weibull percentiles. *Qual. Reliab. Eng. Int.* **2006**, *22*, 141–151. [[CrossRef](#)]
27. Hand, D.J.; Daly, F.; McConway, K.; Lunn, D.; Ostrowski, E. *A Handbook of Small Data Sets*; Chapman & Hall: London, UK, 1994.
28. Chen, Z. Parameter estimation of the Gompertz population. *Biom. J.* **1997**, *39*, 117–124. [[CrossRef](#)]