

On Unified Hybrid Censored Gompertz Distribution: Application to the Wind Speed Data

P. Parviz¹, H. Panahi^{2*} and M. R. Mardanbeigi³

¹Department of Statistics,
Science and Research Branch, Islamic Azad University, Tehran, Iran.
Email: pariyaparviz@gmail.com

²Department of Mathematics and Statistics,
Lahijan Branch, Islamic Azad University, Lahijan, Iran.
Corresponding Author Email: panahi@liau.ac.ir

³Department of Mathematics,
Science and Research Branch, Islamic Azad University, Tehran, Iran.
Email: mrmardanbeigi@srbiau.ac.ir

ABSTRACT

A correct determination of the distribution function is very important in evaluating wind speed because it is a random phenomenon. In this paper, the Gompertz distribution is used to analyze the wind speed data under complete and unified hybrid censored samples. The maximum likelihood estimators of the unknown parameters have been evaluated via an efficient algorithm. The Bayes estimates have been developed using the Metropolis algorithm. Finally, the estimations of the parameters have been studied based on the wind speed data measured in the urban area of Iran. Different probability density functions are employed to describe this data. The quality of the data-fit is distinguished by the several tests and graphical plots. Based on these tests and plots, the Gompertz distribution seems to be the most reliable statistical distribution and thereby make better presentation of the potentials of wind energy.

Keywords: Bayesian method; Gompertz distribution; Maximum likelihood method; Metropolis algorithm; Uniqueness; Unified hybrid censoring; Wind speed; Wind energy.

Mathematics Subject Classification: 62N01, 62N02.

Journal of Economic Literature (JEL) Classification: C11, C13

1. INTRODUCTION

The characterization of the wind speed (WS) is essential for the assessment of wind energy potential. This task is complex due to the random nature of this data. Probability density functions (PDFs) are generally used to describe the WS observations. Several authors considered different distributions for evaluation the WS data. For example, Zhou et al. (2010), Lo Brano (2011), Masseran et al. (2012), Soukissian (2013), Diriba et al. (2017) and Al-Omari and Badran (2009) compared different distributions such as Gamma, Inverse Gamma, Inverse Gaussian, log-Normal, Burr and Erlang for the WS data. In this paper, we propose the Gompertz distribution (Gompertz; 1825) as an alternative

* Corresponding Author: panahi@liau.ac.ir

distribution for characterization the *WS* data in the urban area of Iran. The probability density, the cumulative distribution (CDF) and the reliability functions of the Gompertz can be written as:

$$f(x|\alpha, \beta) = \beta \exp \left[\alpha x - \frac{\beta}{\alpha} (e^{\alpha x} - 1) \right]; \quad x \geq 0, \alpha > 0, \beta > 0, \quad (1)$$

$$F(x|\alpha, \beta) = 1 - \exp \left[-\frac{\beta}{\alpha} (e^{\alpha x} - 1) \right]; \quad x \geq 0, \alpha > 0, \beta > 0,$$

and

$$RF(x|\alpha, \beta) = \exp \left[-\frac{\beta}{\alpha} (e^{\alpha x} - 1) \right].$$

respectively. Moreover in the experimental data, the data are usually censored. Among the different censoring schemes, Type I, Type II and hybrid censoring schemes are the most common censoring schemes (Kundu and Howlader (2010), Patel and Patel (2017) and Mohamed Anmed (2018)). Balakrishnan et al. (2008) proposed an unified hybrid censoring scheme and it can be demonstrated as follows. Fix integers $k, r \in \{0, \dots, n\}$ and $T_1 < T_2 \in (0, \infty)$ such that $k < r$. Now

If the k^{th} failure occurs before time T_1 , the experiment terminate at $\min\{\max(X_{k:n}, T_1), T_2\}$

If the k^{th} failure occurs between T_1 and T_2 , the experiment terminate at $\min\{X_{r:n}, T_2\}$.

If the k^{th} failure occurs after time T_2 , the experiment terminate at Y_k .

Many authors have investigated this censoring scheme for different situations in recent years; see for example, Mohamed (2017), Ateya (2017), Panahi and Sayyereh (2016) and Panahi (2017). Therefore, the following six cases are exist under unified hybrid censoring scheme.

Table 1: The unified hybrid censoring scheme.

Cases	Censoring Scheme	Terminate
I	$0 < X_{k:n} < X_{r:n} < T_1 < T_2$	T_1
II	$0 < X_{k:n} < T_1 < X_{r:n} < T_2$	$X_{r:n}$
III	$0 < X_{k:n} < T_1 < T_2 < X_{r:n}$	T_2
IV	$0 < T_1 < X_{k:n} < X_{r:n} < T_2$	$X_{r:n}$
V	$0 < T_1 < X_{k:n} < T_2 < X_{r:n}$	T_2
VI	$0 < T_1 < T_2 < X_{k:n} < X_{r:n}$	$X_{k:n}$

In this paper, we consider the problems of estimation for the Gompertz distribution under unified hybrid censored sample. The maximum likelihood estimators (MLEs) are calculated by solving one non-linear equation. Based on the Gamma priors for the unknown parameters, the Bayesian estimators have been developed using the Metropolis algorithm. Finally, we consider the Gompertz distribution for characterizing the *WS* data. In Section 2, we compute the MLEs of the unknown parameters using an algorithm. Uniqueness of the MLEs is proved graphically in Section 3. Bayesian inferences are considered in Section 4. The *WS* data are analyzed in Section 5. Conclusions are provided in Section 6.

2. MAXIMUM LIKELIHOOD ESTIMATION

Let $X_i = X_{i:n}; i = 1, \dots, d$ be the unified hybrid censored order statistics from Gompertz distribution. Based on (1), the likelihood function of the parameters α and β can be obtained as

$$l(\alpha, \beta | data) = \begin{cases} \frac{n!}{(n-d)!} \prod_{i=1}^d f(x_{i:n} | \alpha, \beta) [RF(T_1 | \alpha, \beta)]^{n-d}; & d = d_1 = d_2 = r, \dots, n & \text{Case 1} \\ \frac{n!}{(n-r)!} \prod_{i=1}^r f(x_{i:n} | \alpha, \beta) [RF(x_{r:n} | \alpha, \beta)]^{n-r}; & d_1 = k, \dots, r-1, d_2 = r & \text{Case 2} \\ \frac{n!}{(n-d_2)!} \prod_{i=1}^{d_2} f(x_{i:n} | \alpha, \beta) [RF(T_2 | \alpha, \beta)]^{n-d_2}; & d_1 = d_2 = k, \dots, r-1, d_1 \leq d_2 & \text{Case 3} \\ \frac{n!}{(n-r)!} \prod_{i=1}^r f(x_{i:n} | \alpha, \beta) [RF(x_{r:n} | \alpha, \beta)]^{n-r}; & d_1 = k, \dots, r-1, d_2 = r & \text{Case 4} \\ \frac{n!}{(n-d_2)!} \prod_{i=1}^{d_2} f(x_{i:n} | \alpha, \beta) [RF(T_2 | \alpha, \beta)]^{n-d_2}; & d_1 = d_2 = k, \dots, r-1 & \text{Case 5} \\ \frac{n!}{(n-k)!} \prod_{i=1}^k f(x_{i:n} | \alpha, \beta) [RF(x_{k:n} | \alpha, \beta)]^{n-k}; & d_2 = 0, \dots, k-1 & \text{Case 6} \end{cases} \quad (2)$$

where, $d = d_1 = d_2$ and

$$\begin{cases} d_1; & \text{number of failure until } T_1 \\ d_2; & \text{number of failure until } T_2 \end{cases}$$

The logarithm of Equation (2) may then be written as

$$L(\alpha, \beta | data) \propto d \ln \beta + \alpha \sum_{i=1}^d x_{i:n} - \frac{\beta}{\alpha} \sum_{i=1}^d (\exp(\alpha x_{i:n}) - 1) - (n-d) \frac{\beta}{\alpha} (\exp(\alpha v) - 1) \quad (3)$$

Here,

$$(d, v) = \begin{cases} (d, T_1) & \text{Case 1} \\ (r, x_{r:n}) & \text{Case 2 \& 4} \\ (d_2, T_2) & \text{Case 3 \& 5} \\ (k, x_{k:n}) & \text{Case 6} \end{cases}$$

For obtaining the MLEs, we have,

$$\begin{aligned} \frac{\partial L(\alpha, \beta | data)}{\partial \alpha} &= \sum_{i=1}^d x_{i:n} - \beta \frac{\alpha \sum_{i=1}^d x_{i:n} \exp(\alpha x_{i:n}) - \sum_{i=1}^d (\exp(\alpha x_{i:n}) - 1)}{\alpha^2} \\ &\quad - (n-d) \beta \frac{\alpha v \exp(\alpha v) - (\exp(\alpha v) - 1)}{\alpha^2} = 0 \end{aligned}$$

and

$$\frac{\partial L(\alpha, \beta | data)}{\partial \beta} = \frac{d}{\beta} - \frac{1}{\alpha} \sum_{i=1}^d (\exp(\alpha x_{i:n}) - 1) - \frac{n-d}{\alpha} (\exp(\alpha \nu) - 1) = 0 \quad (4)$$

The MLE of the parameter β , say $\hat{\beta}_{MLE}$ is then given, from (4) by

$$\hat{\beta}_{MLE} = \frac{d\alpha}{\sum_{i=1}^d (\exp(\alpha x_{i:n}) - 1) + (n-d)(\exp(\alpha \nu) - 1)}, \quad (5)$$

where $\hat{\alpha}_{MLE}$ can be gained using the solution of the nonlinear equation

$$J(\alpha) = \sum_{i=1}^d x_{i:n} - \frac{d \left(\sum_{i=1}^d x_{i:n} \exp(\alpha x_{i:n}) - (n-d)\nu \exp(\alpha \nu) \right)}{\sum_{i=1}^d (\exp(\alpha x_{i:n}) - 1) + (n-d)(\exp(\alpha \nu) - 1)} + \frac{d}{\alpha} = 0. \quad (6)$$

The solution of Equation (6) can be obtained numerically by some appropriate iterative algorithm. Note that (6) can be considered in the form:

$$\kappa^{-1}(\alpha) = \alpha \quad (7)$$

where

$$\kappa(\alpha) = \frac{1}{d} \left[\frac{d \left(\sum_{i=1}^d x_{i:n} \exp(\alpha x_{i:n}) - (n-d)\nu \exp(\alpha \nu) \right)}{\sum_{i=1}^d (\exp(\alpha x_{i:n}) - 1) + (n-d)(\exp(\alpha \nu) - 1)} - \sum_{i=1}^d x_{i:n} \right].$$

We propose an iterative algorithm to solve (7) as:

Algorithm:

S1- Start with initial value α say $\alpha^{(0)}$.

$$S2 - \text{Compute } \alpha^{(1)} = \kappa^{-1}(\alpha^{(0)}) = \left[\frac{1}{d} \left(\frac{d \left(\sum_{i=1}^d x_{i:n} \exp(\alpha^{(0)} x_{i:n}) - (n-d)\nu \exp(\alpha^{(0)} \nu) \right)}{\sum_{i=1}^d (\exp(\alpha^{(0)} x_{i:n}) - 1) + (n-d)(\exp(\alpha^{(0)} \nu) - 1)} - \sum_{i=1}^d x_{i:n} \right) \right]^{-1}.$$

S3- Calculate $\alpha^{(2)}, \alpha^{(3)}, \dots, \alpha^{(n)}$ similarly.

$$S4 - \text{Obtain } \alpha^{(n+1)} = \kappa(\alpha^{(n)}) = \left[\frac{1}{d} \left(\frac{d \left(\sum_{i=1}^d x_{i:n} \exp(\alpha^{(n)} x_{i:n}) - (n-d)\nu \exp(\alpha^{(n)} \nu) \right)}{\sum_{i=1}^d (\exp(\alpha^{(n)} x_{i:n}) - 1) + (n-d)(\exp(\alpha^{(n)} \nu) - 1)} - \sum_{i=1}^d x_{i:n} \right) \right]^{-1}.$$

S5- Stop when $|\alpha^{(i+1)} - \alpha^{(i)}| < \varepsilon; \quad i = 1, \dots, n.$

$$S6 - \text{Finally, } \hat{\beta}_{MLE} = \frac{d\hat{\alpha}_{MLE}}{\sum_{i=1}^d (\exp(\hat{\alpha}_{MLE} x_{i:n}) - 1) + (n-d)(\exp(\hat{\alpha}_{MLE} v) - 1)}.$$

3. UNIQUENESS OF THE MAXIMUM LIKELIHOOD ESTIMATORS

In this Section, we prove the existence and uniqueness of the MLEs. For this purpose, we introduce the following Steps:

S1- Generate a sample size ($n = 50$) from the Gompertz distribution.

S2- Consider the censoring scheme, $r = 40, k = 15, T_1 = 0.6, T_2 = 1.2.$

S3- Plot $\frac{\partial L(\alpha, \beta | \text{data})}{\partial \beta} = 0$ and $\frac{\partial L(\alpha, \beta | \text{data})}{\partial \alpha} = 0$ in Figure 1.

S4- Plot $l(\alpha, \hat{\beta} | \text{data})$ and $l(\hat{\alpha}, \beta | \text{data})$ in Figures 2 and 3 respectively .

It is found that the MLEs of the parameters α and β exist and unique. Note that the other cases of unified hybrid censoring have the same results.

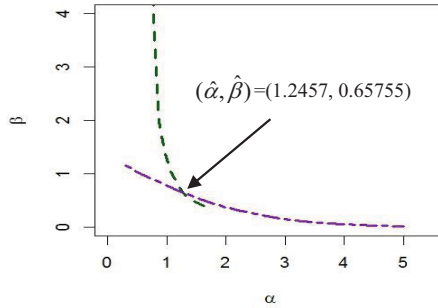


Figure 1: The MLEs of α and β .

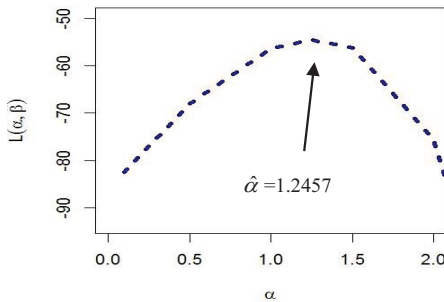


Figure 2: The plot of $L(\alpha, \hat{\beta})$.

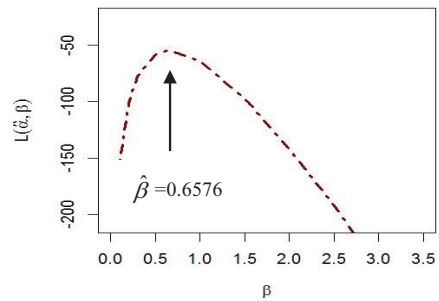


Figure 3: The plot of $L(\hat{\alpha}, \beta)$.

4. BAYESIAN ESTIMATION

In this section, we derive the Bayes estimators of the parameters of the $Gompertz(\alpha, \beta)$. Many authors have discussed Bayesian inference based on different methods, for example, Ali Mousa (2003), Upadhyaya and Gupta (2010) and Panahi (2018). Here, it is assumed that α and β have the following priors:

$$\pi(\alpha) = \text{Gamma}(a_1, b_1),$$

and

$$\pi(\beta) = \text{Gamma}(a_2, b_2).$$

Using the joint prior distribution

$$\pi(\alpha, \beta) \propto \alpha^{a_1-1} e^{-b_1\alpha} \beta^{a_2-1} e^{-b_2\beta}; \quad \alpha, \beta, a_1, b_1, a_2, b_2 > 0 \quad (8)$$

and the likelihood function (2), the joint posterior distribution (JPD) of α and β can be written as:

$$\pi(\alpha, \beta | \text{data}) \propto \begin{cases} \beta^{d+a_2-1} \alpha^{a_1-1} \exp \left[\alpha \sum_{i=1}^d x_{i:n} - \frac{\beta}{\alpha} \sum_{i=1}^d (\exp(\alpha x_{i:n}) - 1) - \frac{\beta(n-d)}{\alpha} (\exp(\alpha T_1) - 1) - b_1\alpha - b_2\beta \right]; & \text{Case 1} \\ \beta^{r+a_2-1} \alpha^{a_1-1} \exp \left[\alpha \sum_{i=1}^r x_{i:n} - \frac{\beta}{\alpha} \sum_{i=1}^r (\exp(\alpha x_{i:n}) - 1) - \frac{\beta(n-r)}{\alpha} (\exp(\alpha x_{r:n}) - 1) - b_1\alpha - b_2\beta \right]; & \text{Case 2} \\ \beta^{d_2+a_2-1} \alpha^{a_1-1} \exp \left[\alpha \sum_{i=1}^{d_2} x_{i:n} - \frac{\beta}{\alpha} \sum_{i=1}^{d_2} (\exp(\alpha x_{i:n}) - 1) - \frac{\beta(n-d_2)}{\alpha} (\exp(\alpha T_2) - 1) - b_1\alpha - b_2\beta \right]; & \text{Case 3} \\ \beta^{r+a_2-1} \alpha^{a_1-1} \exp \left[\alpha \sum_{i=1}^r x_{i:n} - \frac{\beta}{\alpha} \sum_{i=1}^r (\exp(\alpha x_{i:n}) - 1) - \frac{\beta(n-r)}{\alpha} (\exp(\alpha x_{r:n}) - 1) - b_1\alpha - b_2\beta \right]; & \text{Case 4} \\ \beta^{d_2+a_2-1} \alpha^{a_1-1} \exp \left[\alpha \sum_{i=1}^{d_2} x_{i:n} - \frac{\beta}{\alpha} \sum_{i=1}^{d_2} (\exp(\alpha x_{i:n}) - 1) - \frac{\beta(n-d_2)}{\alpha} (\exp(\alpha T_2) - 1) - b_1\alpha - b_2\beta \right]; & \text{Case 5} \\ \beta^{k+a_2-1} \alpha^{a_1-1} \exp \left[\alpha \sum_{i=1}^k x_{i:n} - \frac{\beta}{\alpha} \sum_{i=1}^k (\exp(\alpha x_{i:n}) - 1) - \frac{\beta(n-k)}{\alpha} (\exp(\alpha x_{k:n}) - 1) - b_1\alpha - b_2\beta \right]; & \text{Case 6} \end{cases}$$

Moreover, the JPD of α and β can be combined as:

$$\pi(\alpha, \beta | \text{data}) \propto \beta^{d+a_2-1} \alpha^{a_1-1} \exp \left[\alpha \left(\sum_{i=1}^d x_{i:n} - b_1 \right) - \frac{\beta}{\alpha} \left(\sum_{i=1}^d (\exp(\alpha x_{i:n}) - 1) + (n-d)(\exp(\alpha \nu) - 1) \right) - b_2\beta \right],$$

where, d and ν are defined in Equation (3). Therefore, the Bayes estimators of α and β under the squared error loss function (SEL) are respectively derived as:

$$\hat{\alpha}_{SEL} = \int_0^\infty \int_0^\infty \beta^{d+\alpha_2-1} \alpha^{\alpha_1} \exp \left[\alpha \left(\sum_{i=1}^d x_{i:n} - b_1 \right) - \frac{\beta}{\alpha} \left(\sum_{i=1}^d (\exp(\alpha x_{i:n}) - 1) + (n-d)(\exp(\alpha \nu) - 1) \right) - b_2 \beta \right] d\alpha d\beta ,$$

and

$$\hat{\beta}_{SEL} = \int_0^\infty \int_0^\infty \beta^{d+\alpha_2} \alpha^{\alpha_1-1} \exp \left[\alpha \left(\sum_{i=1}^d x_{i:n} - b_1 \right) - \frac{\beta}{\alpha} \left(\sum_{i=1}^d (\exp(\alpha x_{i:n}) - 1) + (n-d)(\exp(\alpha \nu) - 1) \right) - b_2 \beta \right] d\alpha d\beta .$$

Unfortunately, we cannot compute these estimators analytically. Therefore, we use the following Metropolis algorithm (Metropolis et al.; 1953):

Algorithm:

S1 – Start with $\alpha^0 = \hat{\alpha}_{MLE}$ and $\beta^0 = \hat{\beta}_{MLE}$.

S2 – Put $i = 1$ and then generate

$$\beta^i \sim \text{Gamma}(d + \alpha_2, \omega(\alpha, d, \nu)); \quad \omega(\alpha, d, \nu) = \left[\alpha^{-1} \left(\sum_{i=1}^d (\exp(\alpha x_{i:n}) - 1) + (n-d)(\exp(\alpha \nu) - 1) \right) + b_2 \right]$$

S3 – Consider $\alpha^\dagger \sim \text{Normal}(\alpha^{i-1}, \text{Variance}(\alpha^{i-1}))$

S4 – Obtain the AP as :

$$AP = \min \left[1, \frac{\pi(\alpha^\dagger | \beta^i)}{\pi(\alpha^{i-1} | \beta^i)} \right]; \quad \pi(\alpha | \beta) \propto \alpha^{\alpha_1-1} \exp \left[\alpha \left(\sum_{i=1}^d x_{i:n} - b_1 \right) - \frac{\beta}{\alpha} \left(\sum_{i=1}^d (\exp(\alpha x_{i:n}) - 1) - (n-d)(\exp(\alpha \nu) - 1) \right) \right]$$

$$S5 - \begin{cases} \text{if } Y \leq AP & \text{then } \alpha^i = \alpha^\dagger \\ \text{if } Y > AP & \text{then } \alpha^i = \alpha^{i-1} \end{cases}; Y \sim \text{Uniform}(0, 1).$$

S6 – Put $i = i + 1$ and repeat the Steps N times.

S7 – The Bayesian estimators are :

$$\hat{\alpha}_{SEL} = \frac{1}{N - M} \sum_{i=M+1}^N \alpha^i ,$$

and

$$\hat{\beta}_{SEL} = \frac{1}{N - M} \sum_{i=M+1}^N \beta^i ,$$

where, M is burns.

5. WIND SPEED DATA ANALYSIS

In this section, we illustrate the proposed estimation procedure based on the *WS* data. The data was obtained from the two stations as Ferdous (data set I) and Koohrang (data set II). All stations are installed in the area of Iran. We considered the monthly mean *WS* values from 2008-2010. First, we compared the Gompertz distribution with the other most suitable distributions for the *WS* analysis. These distributions are the log-Normal, Burr XII, Burr III and Inverse Gamma. The MLEs of the parameters for the data sets I and II are presented in Tables 2 and 3 respectively.

Table 2: The MLEs of the unknown parameters.

<i>Data set I:</i>	<i>Parameter Estimates</i>
<i>Gompertz</i>	$\alpha = 1.105901$ and $\beta = 0.002586$
<i>log-Normal</i>	$\mu = 1.614801$ and $\sigma = 0.209409$
<i>Burr XII</i>	$\alpha = 0.049345$ and $\beta = 12.552961$
<i>Burr III</i>	$\alpha = 197.20296$ and $\beta = 3.524023$
<i>Inverse Gamma</i>	$\alpha = 20.56653$ and $\beta = 100.88244$

Table 3: The MLEs of the unknown parameters.

<i>Data set II:</i>	<i>Parameter Estimates</i>
<i>Gompertz</i>	$\alpha = 0.517874$ and $\beta = 0.074099$
<i>log-Normal</i>	$\mu = 1.106842$ and $\sigma = 0.519642$
<i>Burr XII</i>	$\alpha = 0.0176012$ and $\beta = 51.274886$
<i>Burr III</i>	$\alpha = 6.222710$ and $\beta = 2.028057$
<i>Inverse Gamma</i>	$\alpha = 3.58080$ and $\beta = 9.360822$

The two criteria such as Akaike information criterion (AIC) and Bayesian information criterion (BIC) are considered. Based on data set I, the (AIC, BIC) for the Gompertz, log-Normal, Burr XII, Burr III and Inverse Gamma are (99.697691, 102.86472), (109.868, 113.0351), (226.7683, 229.9353), (131.2011, 134.3682) and (114.1627, 117.3297) respectively. Also for data set II the (AIC, BIC) are (136.25309, 139.42013), (138.7228, 141.8898), (165.8297, 168.9968), (146.3442, 149.5112) and (143.2654, 146.4324) respectively. It is observed that the Gompertz distribution has the smallest AIC and BIC values. The goodness-of-fit statistics for data sets I and II are given by:

- **Gompertz**

Kolmogorov - Smirnov statistic : 0.0933710 (data set I) and 0.097274 (data set II)
Cramer - von Mises statistic : 0.0814324 (data set I) and 0.086598 (data set II)
Anderson - Darling statistic : 0.5987421 (data set I) and 0.610326 (data set II)

- **log-Normal**

Kolmogorov - Smirnov statistic : 0.1286736 (data set I) and 0.1264787 (data set II)
Cramer - von Mises statistic : 0.1092808 (data set I) and 0.1185824 (data set II)

Anderson – Darling statistic : 0.7815657 (data set I) and 0.7539940 (data set II)

• **Burr XII**

Kolmogorov – Smirnov statistic : 0.5070617 (data set I) and 0.3000545 (data set II)

Cramer – von Mises statistic : 2.7500361 (data set I) and 0.8949390 (data set II)

Anderson – Darling statistic : 12.811449 (data set I) and 4.3982313 (data set II)

• **Burr III**

Kolmogorov – Smirnov statistic : 0.2250839 (data set I) and 0.1551290 (data set II)

Cramer – von Mises statistic : 0.4334767 (data set I) and 0.2092584 (data set II)

Anderson – Darling statistic : 2.7201088 (data set I) and 1.2547253 (data set II)

• **Inverse Gamma**

Kolmogorov – Smirnov statistic : 0.1440171 (data set I) and 0.1539733 (data set II)

Cramer – von Mises statistic : 0.1576876 (data set I) and 0.2021529 (data set II)

Anderson – Darling statistic : 1.0831201 (data set I) and 1.1863734 (data set II)

Also, the p -values of the data sets I and II are 0.9122 and 0.8852 respectively. Based on the p -values and the goodness-of-fit statistics, we can say that the Gompertz distribution fits quite fine to the above data more than any other distribution. We present the probability-probability (P-P) plots of the above five distributions in Figures 4 and 5 respectively.

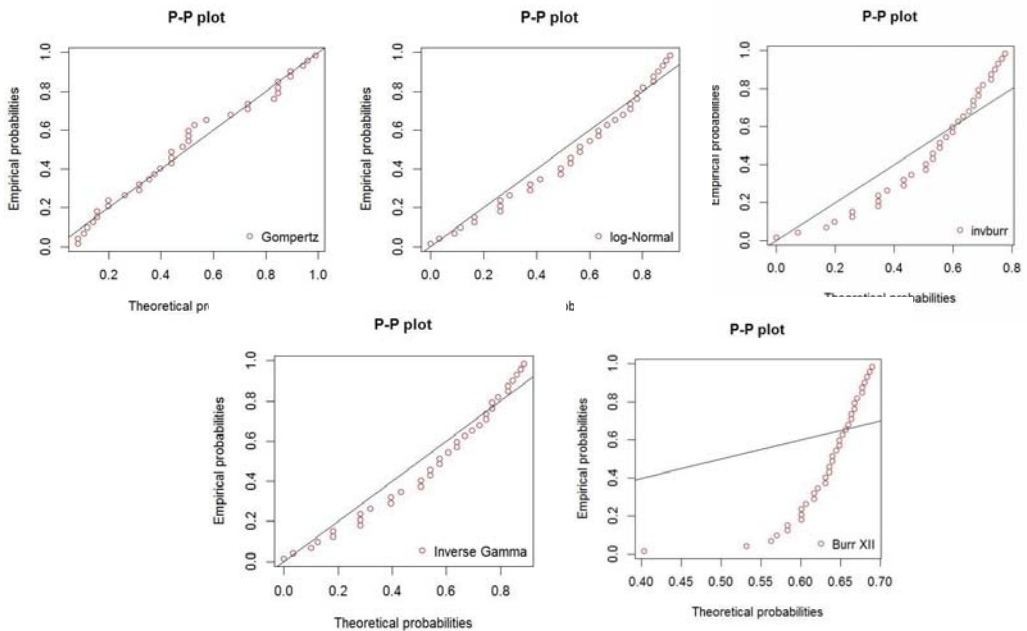


Figure 4: The P-P plots for data set I.

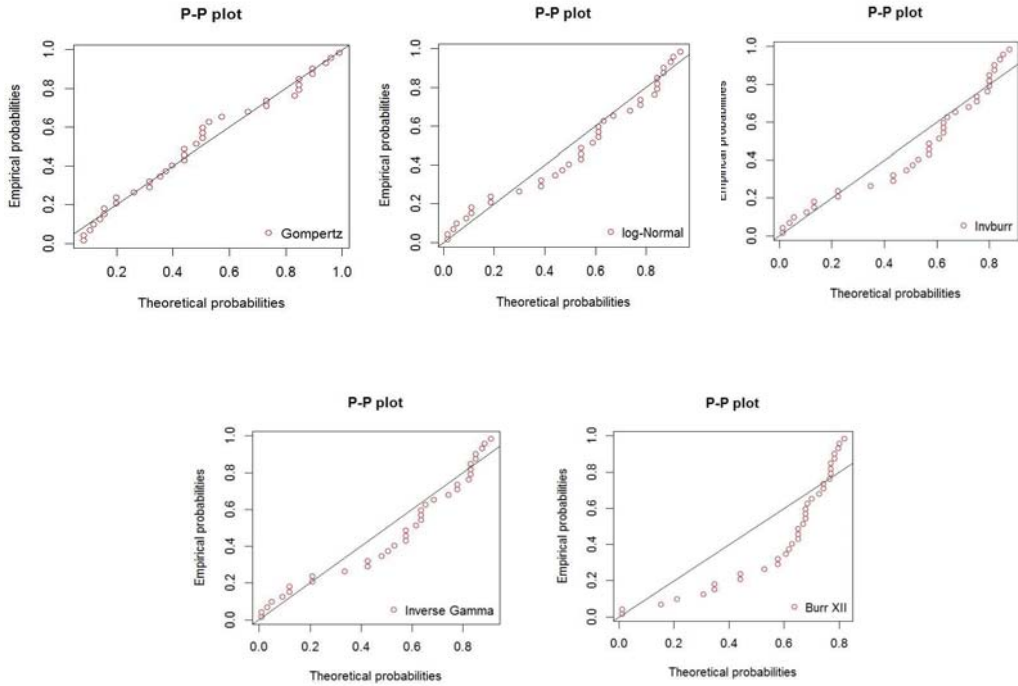


Figure 5: The P-P plots for data set II.

Therefore, the Gompertz distribution is a better model than the other proposed distributions. Based on the Gompertz distribution, we also plot the empirical survival function, and the fitted survival function in Figures 6 and 7. Observe that the fitted Gompertz survival function is almost idem, and it fits the data very good.

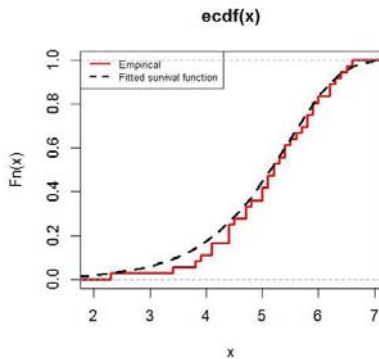


Figure 6: Empirical and fitted survival function for the data set I.

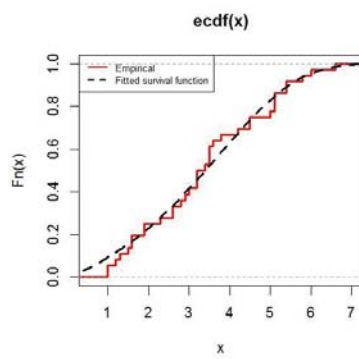


Figure 7: Empirical and fitted survival function for the data set II.

Now, we propose different censoring schemes of the two data sets for illustrative purposes. We generate unified hybrid censored data from these data sets and report them in the Tables 4 and 5 respectively.

Table 4: The unified hybrid censoring scheme for data set I.

Case	Censoring Scheme	Terminate
I	$X_{k:n} = 3.4, X_{r:n} = 4.5, T_1 = 5.0, T_2 = 6.0$	T_1
II	$X_{k:n} = 3.4, X_{r:n} = 5.5, T_1 = 4.6, T_2 = 6.0$	$X_{r:n}$
III	$X_{k:n} = 3.4, X_{r:n} = 6.4, T_1 = 4.6, T_2 = 5.9$	T_2
IV	$X_{k:n} = 5.1, X_{r:n} = 6.2, T_1 = 4.6, T_2 = 6.4$	$X_{r:n}$
V	$X_{k:n} = 5.1, X_{r:n} = 6.5, T_1 = 4.6, T_2 = 6.1$	T_2
VI	$X_{k:n} = 5.8, X_{r:n} = 6.5, T_1 = 4.6, T_2 = 5.4$	$X_{k:n}$

Table 5: The unified hybrid censoring scheme for data set II.

Case	Censoring Scheme	Terminate
I	$X_{k:n} = 1.2, X_{r:n} = 2.3, T_1 = 3.9, T_2 = 5.2$	T_1
II	$X_{k:n} = 1.2, X_{r:n} = 4.2, T_1 = 3.9, T_2 = 5.2$	$X_{r:n}$
III	$X_{k:n} = 1.2, X_{r:n} = 5.7, T_1 = 3.9, T_2 = 5.2$	T_2
IV	$X_{k:n} = 3.2, X_{r:n} = 5.4, T_1 = 2.8, T_2 = 5.9$	$X_{r:n}$
V	$X_{k:n} = 3.2, X_{r:n} = 5.4, T_1 = 2.8, T_2 = 4.7$	T_2
VI	$X_{k:n} = 5.1, X_{r:n} = 5.9, T_1 = 2.8, T_2 = 4.7$	$X_{k:n}$

For all the six Cases of the two data sets, we assess the unknown parameters using the MLEs and the Bayesian estimates. To gain the MLE of α , we have utilized the iterative algorithm which is described in Section 2 and stopped the iterative procedure when the difference between two successive iterates is less than 10^{-6} . To compute the Bayesian estimates, we require specifying prior distributions of the parameters. We have considered the non-informative ($a_1 = b_1 = a_2 = b_2 = 0$) prior. The Bayesian estimates are obtained using the maximum likelihood estimates of α and β as an initial guess for running the proposed Bayesian algorithm. For two data sets the maximum likelihood estimates and the Bayesian estimates are reported in Tables 6 and 7 respectively.

Table 6: Different estimates of the parameters α and β for data set I.

<i>I</i>	<i>II</i>	<i>III</i>	<i>IV</i>	<i>V</i>	<i>VI</i>
MLE					
$\alpha = 1.69026$	$\alpha = 1.57502$	$\alpha = 1.43193$	$\alpha = 1.35714$	$\alpha = 1.37031$	$\alpha = 1.44903$
$\beta = 1.97 \times 10^{-3}$	$\beta = 2.05 \times 10^{-3}$	$\beta = 2.20 \times 10^{-3}$	$\beta = 2.31 \times 10^{-3}$	$\beta = 2.29 \times 10^{-3}$	$\beta = 2.19 \times 10^{-3}$
Bayes					
$\alpha = 1.67824$	$\alpha = 1.55743$	$\alpha = 1.41023$	$\alpha = 1.34025$	$\alpha = 1.35963$	$\alpha = 1.42989$
$\beta = 1.96 \times 10^{-3}$	$\beta = 2.03 \times 10^{-3}$	$\beta = 2.19 \times 10^{-3}$	$\beta = 2.29 \times 10^{-3}$	$\beta = 2.28 \times 10^{-3}$	$\beta = 2.17 \times 10^{-3}$

Table 7: Different estimates of the parameters α and β for data set II.

<i>I</i>	<i>II</i>	<i>III</i>	<i>IV</i>	<i>V</i>	<i>VI</i>
MLE					
$\alpha = 0.93073$	$\alpha = 0.81687$	$\alpha = 0.71578$	$\alpha = 0.68037$	$\alpha = 0.75221$	$\alpha = 0.73451$
$\beta = 0.05879$	$\beta = 0.06171$	$\beta = 0.06691$	$\beta = 0.06864$	$\beta = 0.06381$	$\beta = 0.06512$
Bayes					
$\alpha = 0.90995$	$\alpha = 0.79685$	$\alpha = 0.69734$	$\alpha = 0.66542$	$\alpha = 0.73012$	$\alpha = 0.72115$
$\beta = 0.05626$	$\beta = 0.06043$	$\beta = 0.06515$	$\beta = 0.06706$	$\beta = 0.06258$	$\beta = 0.06402$

6. CONCLUSIONS

Selecting a correct distribution is important in *WS* data analysis. In this article, we analyzed the complete and censored *WS* data using the Gompertz distribution. We calculated the MLEs of the unknown parameters using an efficient algorithm. We also provided the Bayes estimators of the parameters α and β by applying the Metropolis algorithm. Based on the results for all distributions and with respect to all goodness of fit and graphical methods, the Gompertz distribution was found to be the most appropriate for describing the *WS* data measured in the urban area of Iran. Thus, the Gompertz can be utilized as an alternative distribution that effectively explains *WS* data. Moreover, the proposed estimators work well for estimating the unknown parameters under censored data.

APPENDIX:

The knowledge of the statistical distribution of the *WS* is very important. So in this Appendix, the statistical properties (*SP*) of the four used distributions are described.

• Burr XII distribution

The Burr distributions was proposed by Burr (Burr; 1942) for the represented purpose of fitting a CDF to a variety of frequency data forms. Table 8 illustrated the *SP* of the Burr XII distribution, in which α and β are the shape parameters.

Table 8: The *SP* of the Burr XII distribution.

Burr XII	
PDF	$f(x; \alpha, \beta) = \alpha\beta x^{\beta-1} (1+x^\beta)^{-\alpha-1}; x > 0, \alpha, \beta > 0$
CDF	$F(x; \alpha, \beta) = 1 - (1+x^\beta)^{-\alpha}; x > 0, \alpha, \beta > 0$
Mean	$\Gamma(\alpha - \frac{1}{\beta})\Gamma(1 + \frac{1}{\beta}) / \Gamma(\alpha)$
Variance	$\left(\Gamma(\alpha - \frac{2}{\beta})\Gamma(1 + \frac{2}{\beta}) - \left(\Gamma(\alpha - \frac{1}{\beta})\Gamma(1 + \frac{1}{\beta}) \right)^2 / \Gamma(\alpha) \right) / \Gamma(\alpha)$

• Burr III distribution

If the random variable *Y* has a Burr XII distribution, then the random variable $X = 1/Y$ has a Burr III distribution. This distribution has been utilized widely in abundant fields of sciences with different parameterizations. The properties of Burr III distribution are presented in Table 9.

Table 9: The *SP* of the Burr III distribution.

Burr III	
PDF	$f(x; \alpha, \beta) = \alpha\beta x^{-\beta-1} (1+x^{-\beta})^{-\alpha-1}; x > 0, \alpha, \beta > 0$
CDF	$F(x; \alpha, \beta) = (1+x^\beta)^{-\alpha}; x > 0, \alpha, \beta > 0$
Mean	$\left(-\Gamma(-\frac{1}{\beta})\Gamma(\alpha + \frac{1}{\beta}) \right) / \beta\Gamma(\alpha); \beta > 1$
Variance	$\left(-2\beta\Gamma(-\frac{2}{\beta})\Gamma(\alpha + \frac{2}{\beta}) - \left(\Gamma(-\frac{1}{\beta})\Gamma(\alpha + \frac{1}{\beta}) \right)^2 / \Gamma(\alpha) \right) / \beta^2\Gamma(\alpha)$

• log-Normal distribution

The log-Normal distribution is one of the most widely used statistical distributions in different sciences.(Zhu et al. (2012)) Table 10 represented the *SP* of the log-Normal distribution.

Table 10: The *SP* of the log-Normal distribution.

log-Normal	
PDF	$f(x; \mu, \sigma) = \frac{1}{x\sigma\sqrt{2\pi}} e^{-\frac{(\ln x - \mu)^2}{2\sigma^2}}; x > 0, \mu \in (-\infty, +\infty), \sigma > 0$
CDF	$F(x; \mu, \sigma) = \Phi\left(\frac{\ln x - \mu}{\sigma}\right); x > 0, \mu \in (-\infty, +\infty), \sigma > 0$
Mean	$e^{\mu + \sigma^2/2}$
Variance	$e^{2\mu + \sigma^2} (e^{\sigma^2} - 1)$

- **Inverse Gamma distribution**

Table 11 illustrated the *SP* of the Inverse Gamma distribution in which α is the shape parameter, β is the scale parameter and Γ is the Gamma Function.

Table 11: The *SP* of the Inverse Gamma distribution.

PDF	$f(x; \alpha, \beta) = \left(\beta^\alpha / \Gamma(\alpha)\right) x^{-\alpha-1} e^{-\beta/x}; x > 0, \alpha, \beta > 0$
CDF	$F(x; \alpha, \beta) = \Gamma(\alpha, \beta/x) / \Gamma(\alpha); x > 0, \alpha, \beta > 0$
Mean	$\frac{\beta}{\alpha - 1}; \alpha > 1$
Variance	$\frac{\beta^2}{(\alpha - 1)^2 (\alpha - 2)}; \alpha > 2$

7. REFERENCES

- Ali Mousa, M.A.M., 2003, Bayesian prediction based on Pareto doubly censored data. *Statistics*. **37**,65-72.
- Al-Omari, J.F. M., Badran, O.O., 2009, Modeling of Weather Characteristics in Jordan, *Int. J. Ecology. Econ. Stat.* **15**, 93-102.
- Ateya, S.F., 2017, Estimation under inverse Weibull distribution based on Balakrishnan's unified hybrid censored scheme. *Commun. Stat. Simul. Comput.* **46**,3645-3666.
- Balakrishnan, N., Rasouli, A., Sanjari Farsipour, N., 2008, Exact likelihood inference based on an unified hybrid censored sample from the exponential distribution. *J. Stat. Comput. Simul.* **78**, 475-488.
- Burr, I.W., 1942, Cumulative frequency functions. *Ann. Math. Stat.* **13**, 215–232.
- Diriba, T.A., Debusho, L.K., Botai, J., Hassen, A., 2017, Bayesian modelling of extreme wind speed at Cape Town, South Africa. *Environ. Ecol. Stat.* **24**,243-267.

- Gompertz, B., 1825, On the nature of the function expressive of the law of human mortality and on a new model of determining the value of life contingencies. *Philos. Trans. R. Soc. Lond.* **115**,513–585.
- Kundu, D., Howlader, H., 2010, Bayesian inference and prediction of the inverse Weibull distribution for Type-II censored data. *Comput. Stat. Data Analys.* **54**, 1547-1558.
- Lo Brano, V., Orioli, A., Ciulla, G., Culotta, S., 2011, Quality of wind speed fitting distributions for the urban area of Palermo, Italy. *Renew Energy*. **36**,1026-1039.
- Masseran, N., Razali, A.M., Ibrahim, K., 2012, An analysis of wind power density derived from several wind speed density functions: the regional assessment on wind power in Malaysia. *Renew Sustain. Energy Rev.* **16**,6476-6487.
- Metropolis, N., Rosenbluth, A.W., Rosenbluth, M.N., Teller, A.H., Teller, E., 1953, Equations of state calculations by fast computing machines. *J. Phys. Chem.* **21**,, 1087-1092.
- Mohamed H., 2017, Exponentiated Rayleigh Distribution: A Bayes Study Using MCMC Approach Based on Unified Hybrid Censored Data. *J. Advances Math.* **12**, 6863-6880.
- Mohamed Anmed, A.O., 2018, Bayesian Using Extension Jeffreys Prior for Weibull Regression Censored Data. *Int. J. Appl. Math. Stat.* **57**, 74-83.
- Patel, D.T., Patel, M.N., 2017, A Bayesian Approach to Optimal Warranty Length for an Exponential Life Time Distributed Product with Multiply Type-II Censoring Scheme. *Int. J. Math. Stat.* **18**,47-63.
- Panahi, H., Sayyareh, A., 2016, Estimation and prediction for a unified hybrid-censored Burr Type XII distribution. *J. Stat. Comput. Simul.* **86**, 55-73.
- Panahi, H., 2017, Estimation of the Burr type III distribution with application in unified hybrid censored sample of fracture toughness. *J. App. Statis.* **44**,2575-2592.
- Panahi, H., 2018, Estimation for the parameters of the Burr Type XII distribution under doubly censored sample with application to microfluidics data. *International Journal of System Assurance Engineering and Management.* 1-19.
- Soukissian, T., 2013, Use of multi-parameter distributions for offshore wind speed modeling: the Johnson SB distribution. *Appl. Energy*. **111**, 982-1000.
- Upadhyaya, S.K., Gupta, A., 2010, A Bayes analysis of modified Weibull distribution via Markov chain Monte Carlo simulation. *J. Stat. Comput. Simul.* **80**,241–254.
- Zhu, Z., Zhou, J., Chen, L., 2012, Application of Log-Normal Distribution and Data Mining Method in Component Repair Time Calculation of Power System Operation Risk Assessment. *Adv. Electr. Comp. Eng.* **139**,423-431.
- Zhou, J.Y., Erdem, E., Li, G., Shi, J., 2010, Comprehensive evaluation of wind speed distribution models: a case study for North Dakota sites. *Energy Convers. Manage.* **51**,1449-1458.