Prediction of future failures for generalized exponential distribution under Type-I or Type-II hybrid censoring

R. Valiollahi[a], **A. Asgharzadeh**[b*], **D. Kundu**[c]

[a]*Department of Mathematics, Statistics and Computer Science, Semnan University, Semnan, Iran*

[b]*Department of Statistics, Faculty of Mathematical Sciences, University of Mazandaran, Babolsar, Iran*

[c]*Department of Mathematics and Statistics, Indian Institute of Technology, Kanpur, India*

Abstract

In this paper, we consider the prediction of a future observation based on either Type-I or Type-II hybrid censored samples when the lifetime distribution of the experimental units is assumed to be a generalized exponential random variable. Different point and interval predictors are obtained using classical and Bayesian approaches. Monte Carlo simulations are performed to compare the performances of the different methods, and the analysis of one data set has been presented for illustrative purposes.

Keywords: Type-I hybrid censoring, Type-II hybrid censoring, Generalized exponential distribution, Predictor, Prediction interval, Monte Carlo simulation.

1 Introduction

The generalized exponential (GE) distribution is one of the most widely used distributions in reliability and survival analysis. It has been used quite successfully to analyze lifetime data in place of Weibull or gamma distribution. Because of its explicit expression of the cumulative distribution function (CDF), it has been used quite conveniently for analyzing censored data. The probability density function (PDF) and hazard function can take variety of shapes. The PDF of the GE distribution can be a decreasing or an unimodal function, and the hazard function can be either increasing, decreasing or constant, depending on the shape parameter.

*Corresponding author:

E-mail address: a.asgharzadeh@umz.ac.ir

In this manuscript it is assumed that the two-parameter generalized exponential distribution has the following PDF and CDF

$$f(x;\alpha,\lambda) = \begin{cases} \alpha\lambda e^{-\lambda x}\left(1-e^{-\lambda x}\right)^{\alpha-1} & \text{if} \quad x>0 \\ 0 & \text{otherwise} \end{cases} \tag{1.1}$$

and

$$F(x;\alpha,\lambda) = \begin{cases} \left(1-e^{-\lambda x}\right)^{\alpha} & \text{if} \quad x>0 \\ 0 & \text{otherwise,} \end{cases} \tag{1.2}$$

respectively. Here $\alpha > 0$ is the shape parameter and $\lambda > 0$ is the scale parameter. From now on, a two-parameter generalized exponential distribution with the PDF (1.1) and CDF (1.2) will be denoted by GE(α, λ).

Censoring is very common in any reliability or lifetesting experiment. The complete survival times may not be observed by the experimenter most of times. The two most common censoring schemes are namely Type-I and Type-II censoring. They can be briefly described as follows. Consider a sample of n units is placed on a life testing experiment at the time point zero. In Type-I censoring scheme, a time T, independent of the failure times, is pre-fixed so that beyond this time no failures will be observed, that is, the experiment terminates at the time point T. In Type-II censoring scheme, the number of observed failures is fixed, say r ($r \leq n$), and the experiment stops when the r-th failure takes place. Epstein (1954) introduced a new censoring scheme which is a mixture of Type-I and Type-II censoring schemes, and called it as the hybrid censoring scheme. From now on, we will call it as the Type-I hybrid censoring scheme (Type-I HCS). Type-I hybrid censoring scheme can be described as follows. Suppose n items are put on a life testing experiment at the time point zero. Let $X_{1:n} < \ldots < X_{n:n}$ denote the ordered lifetime of the experimental units. In Type-I HCS, the experiment stops at $T_1^* = \min\{X_{r:n}, T\}$, where r and T are pre-fixed.

The main advantage of Type-I HCS is that the experiment will not last more than T units of time, and also more than r units will not fail. The major disadvantage of a Type-I HCS is that there is a possibility that very few failures may occur before time T. Hence the inference of the lifetime distribution of the experimental units will not be very accurate. For this reason, Childs et al. (2003) proposed a new censoring scheme, known as the Type-II hybrid censoring scheme (Type-II HCS), which guarantees a fixed number of failures. In this case, the termination point is $T_2^* = \max\{X_{r:n}, T\}$. For a detailed comparison of Type-I and Type-II HCS, the readers are referred to Childs et al. (2003).

Recently, the hybrid censoring scheme has received considerable interest among the statisticians. Epstein (1954) first introduced Type-I HCS and analyzed the data under the assumption of exponential lifetime distribution of the experimental units. Since the introduction of Type-I HCS by Epstein (1954), extensive work has been done on hybrid censoring and many different variations of it. See, for example, Fairbanks et al. (1982), Draper and Guttman (1987), Chen and Bhattacharya (1988), Ebrahimi (1986 , 1992), Jeong et al. (1996), Kundu and Gupta (1988), Childs et al. (2003), Kundu (2007), Kundu and Banerjee (2008) and the recent review article by Balakrishnan and Kundu (2012) on this topic.

One of the most important problems in life testing is predicting future failures given a record of observed failures. Information regarding future observations can tell us at an early stage of testing how costly the testing is and whether actions should be taken to redesign the test. Extensive work has been done on prediction problem based on frequentist and Bayesian framework. Kaminsky and Rhodin (1985) have applied maximum likelihood to the joint prediction of a future random variable and unknown parameter. Smith (1997, 1999) discussed the properties of the different predictors based on Bayes and frequentist procedures for a class of parametric family of distribution functions, under smooth loss functions. Dellaportas and Wright (1991) used a numerical method to Bayesian prediction for the two-parameter Generalized exponential distribution under the squared error loss function. Raqab et al. (2010), Balakrishnan et al. (2010) and Basak et al. (2006) have discussed different predictors of times to failure of units censored in progressively Type-II censored samples. Kundu and Raqab (2012) considered the prediction of future observation from a Type-II censored data for the two-parameter Weibull distribution under fairly flexible priors on the shape and scale parameters. Recently, Asgharzadeh et al. (2013) discussed the prediction of future observation based on a Type-I hybrid censored data from the two-parameter Weibull distribution.

The main aim of this paper is to consider the prediction of future observation based on Type-I or Type-II hybrid censored samples from a two-parameter Generalized exponential distribution. We obtain the maximum likelihood predictor, best unbiased predictor, conditional median predictor and Bayesian predictor. We also obtain different prediction intervals. It is difficult to compare the performances of the different methods theoretically. We perform some Monte Carlo simulations to compare the performances of the different methods. Finally one data analysis has been performed for illustrative purposes.

Rest of the paper is organized as follows. In Section 2, we provide the preliminaries and notations. Different classical predictors are presented in Section 3. In Section 4, we discuss Bayesian prediction of future observations. In Section 5, one data analysis has been presented. Monte Carlo simulation results are presented in Section 6. Finally we conclude the paper in Section 6.

2 Notation and preliminaries

Consider $\mathbf{X} = (X_{1:n}, X_{2:n}, \ldots, X_{D:n})$ as a sample from model (1.1) which is obtained based on a Type-I HCS or Type-II HCS. Here D is the number of failures during the experiment. In Type-I HCS, $D = r$ if $X_{r:n} \leq T$, and $D = J < r$ if $X_{r:n} > T$. For Type-II HCS, $D = r$ if $X_{r:n} \geq T$, and $D = J \geq r$ if $X_{r:n} < T$. In both the cases, J is determined such that $X_{J:n} < T < X_{J+1:n}$ where $J = 0, \ldots, n$.

For notational simplicity, we will write $(X_1, X_2, \ldots, X_D)$ for $(X_{1:n}, X_{2:n}, \ldots, X_{D:n})$. Based on the observed data, the likelihood function for α and λ without the normalizing constant is

$$L(\alpha, \lambda) = \alpha^D \lambda^D e^{-\lambda \sum_{i=1}^{D} x_i} \left[1 - \left(1 - e^{-\lambda T_0}\right)^{\alpha}\right]^{n-D} \left\{\prod_{i=1}^{D} \left(1 - e^{-\lambda x_i}\right)^{\alpha - 1}\right\}, \tag{2.1}$$

where T_0 denotes the time when the experiment is stopped. In other words, for Type-I HCS, $T_0 = X_{r:n}$ if $X_{r:n} \leq T$ and $T_0 = T$ if $X_{r:n} > T$. For Type-II HCS, $T_0 = X_{r:n}$ if $X_{r:n} \geq T$ and $T_0 = T$ if $X_{r:n} < T$.

From (2.1), the maximum likelihood estimators (MLEs) of α and λ can be obtained as solutions of the following equations

$$\frac{D}{\alpha} + \sum_{i=1}^{D} \ln\left(1 - e^{-\lambda x_i}\right) - (n - D)\frac{\left(1 - e^{-\lambda T_0}\right)^{\alpha} \ln\left(1 - e^{-\lambda T_0}\right)}{1 - \left(1 - e^{-\lambda T_0}\right)^{\alpha}} = 0, \tag{2.2}$$

and

$$\frac{D}{\lambda} - \sum_{i=1}^{D} x_i + (\alpha - 1)\frac{x_i e^{-\lambda x_i}}{1 - e^{-\lambda x_i}} - (n - D)\frac{\alpha T_0 e^{-\lambda T_0}\left(1 - e^{-\lambda T_0}\right)^{\alpha - 1}}{1 - \left(1 - e^{-\lambda T_0}\right)^{\alpha}}. \tag{2.3}$$

The main aim of this paper is to discuss different methods of prediction of $Y = X_{s+D:n} (s = 1, 2, \ldots, n - D)$ of all the $n - D$ censored units based on observed data $\mathbf{X} = (X_1, \ldots, X_D)$. Due to the Markovian property of censored-order statistics, the conditional distribution of Y given $\mathbf{X} = \mathbf{x}$ is just the distribution of $X_{s+D:n}$ given $X_D = x_D$. This implies that the density of Y given $\mathbf{X} = \mathbf{x}$ is the same as the density of the sth order statistic out of $n - D$ units from the population with density $f(y)/(1 - F(T_0)), y \geq T_0$ (left truncated density at T_0). Therefore, the conditional density of $Y = X_{s+D:n}$ given $\mathbf{X} = \mathbf{x}$, for $y \geq T_0$, is given by

$$f(y|\mathbf{x}) = s\binom{n - D}{s} f(y) \left[F(y) - F(T_0)\right]^{s-1} \left[1 - F(y)\right]^{n-D-s} \left[1 - F(T_0)\right]^{-(n-D)}, \tag{2.4}$$

where $s = 1, 2, \ldots, n - D$. For model (1.1), (2.4) reduces to

$$\begin{aligned} f(y|\mathbf{x}, \alpha, \lambda) &= s\binom{n-D}{s}\alpha\lambda e^{-\lambda y}\left(1-e^{-\lambda y}\right)^{\alpha-1}\left[\left(1-e^{-\lambda y}\right)^{\alpha}-\left(1-e^{-\lambda T_0}\right)^{\alpha}\right]^{s-1} \\ &\times \left[1-\left(1-e^{-\lambda y}\right)^{\alpha}\right]^{n-D-s}\left[1-\left(1-e^{-\lambda T_0}\right)^{\alpha}\right]^{-(n-D)}, \end{aligned} \tag{2.5}$$

which is the conditional density of $Y = X_{s+D:n}$ given $\mathbf{X} = (X_1, ..., X_D)$.

3 Classical prediction of future observations

In this section, we obtain several predictors of $Y = X_{s+D:n}$ on the basis of $\mathbf{X} = (X_1, ..., X_D)$ via classical approaches.

3.1 Likelihood prediction approach

In likelihood prediction approach, the principle of maximum likelihood is applied to the joint prediction and estimation of a future random variable and unknown parameters. We assume dependence between present and future, and the approach is non-Bayesian. The predictive likelihood function (PLF) of Y and (α, λ) is given by

$$L(y, \alpha, \lambda|\mathbf{x}) = f(y|\mathbf{x}, \alpha, \lambda)f(\mathbf{x}|\alpha, \lambda). \tag{3.1}$$

Suppose $\widehat{Y} = u(\mathbf{X})$, $\widehat{\alpha} = v_1(\mathbf{X})$ and $\widehat{\lambda} = v_2(\mathbf{X})$ are statistics for which

$$L(u(\mathbf{x}), v_1(\mathbf{x}), v_2(\mathbf{x})|\mathbf{x}) = \sup_{(y,\alpha,\lambda)} L(y, \alpha, \lambda|\mathbf{x}),$$

then $u(\mathbf{X})$ is said to be the maximum likelihood predictor (MLP) of Y and $v_1(\mathbf{x})$ and $v_2(\mathbf{x})$ the predictive maximum likelihood estimators (PMLEs) of α and λ, respectively.

For the GE model, using (2.1) and (2.5), the predictive likelihood function (PLF) of Y, α and λ, is given by

$$\begin{aligned} L(y, \alpha, \lambda) &= s\binom{n-D}{s}(\alpha\lambda)^{d+1}\left(1-e^{-\lambda y}\right)^{\alpha-1}\left\{\prod_{i=1}^{D}\left(1-e^{-\lambda x_i}\right)^{\alpha-1}\right\} \\ &\times e^{-\lambda[y+\sum_{i=1}^{D}x_i]}\left[\left(1-e^{-\lambda y}\right)^{\alpha}-\left(1-e^{-\lambda T_0}\right)^{\alpha}\right]^{s-1}\left[1-\left(1-e^{-\lambda y}\right)^{\alpha}\right]^{n-D-s}. \end{aligned} \tag{3.2}$$

Apart from a constant term, the predictive log-likelihood function is

$$\begin{aligned}\ln L(y,\alpha,\lambda) &= (D+1)[\ln(\alpha)+\ln(\lambda)] + (\alpha-1)\left[\ln[1-e^{-\lambda y}] + \sum_{i=1}^{D}\ln[1-e^{-\lambda x_i}]\right] \\ &- \lambda\left[y+\sum_{i=1}^{D}x_i\right] + (s-1)\ln\left[\left(1-e^{-\lambda y}\right)^{\alpha} - \left(1-e^{-\lambda T_0}\right)^{\alpha}\right] \\ &+ (n-D-s)\ln[1-\left(1-e^{-\lambda y}\right)^{\alpha}]. \end{aligned} \tag{3.3}$$

By using (3.3), the predictive likelihood equations (PLEs) for y, α and λ are given, respectively, by

$$\begin{aligned}\frac{\partial \ln L(y,\alpha,\lambda)}{\partial y} &= -\lambda + (\alpha-1)\frac{\lambda e^{-\lambda y}}{1-e^{-\lambda y}} + (s-1)\frac{\alpha\lambda e^{-\lambda y}\left(1-e^{-\lambda y}\right)^{\alpha-1}}{(1-e^{-\lambda y})^{\alpha}-(1-e^{-\lambda T_0})^{\alpha}} \\ &- (n-D-s)\frac{\alpha\lambda e^{-\lambda y}\left(1-e^{-\lambda y}\right)^{\alpha-1}}{1-(1-e^{-\lambda y})^{\alpha}} = 0, \end{aligned} \tag{3.4}$$

$$\begin{aligned}\frac{\partial \ln L(y,\alpha,\lambda)}{\partial \alpha} &= \ln[1-e^{-\lambda y}] + \sum_{i=1}^{D}\ln[1-e^{-\lambda x_i}] - (n-D-s)\frac{\left(1-e^{-\lambda y}\right)^{\alpha}\ln[1-e^{-\lambda y}]}{1-(1-e^{-\lambda y})^{\alpha}} \\ &+\frac{D+1}{\alpha} + (s-1)\frac{\left(1-e^{-\lambda y}\right)^{\alpha}\ln[1-e^{-\lambda y}] - \left(1-e^{-\lambda T_0}\right)^{\alpha}\ln[1-e^{-\lambda T_0}]}{(1-e^{-\lambda y})^{\alpha}-(1-e^{-\lambda T_0})^{\alpha}} = 0, \end{aligned} \tag{3.5}$$

$$\begin{aligned}\frac{\partial \ln L(y,\alpha,\lambda)}{\partial \lambda} &= \frac{D+1}{\lambda} - [y+\sum_{i=1}^{D}x_i] + (\alpha-1)\left[\frac{ye^{-\lambda y}}{1-e^{-\lambda y}} + \sum_{i=1}^{D}\frac{x_i e^{-\lambda x_i}}{1-e^{-\lambda x_i}}\right] \\ &+(s-1)\frac{\alpha y e^{-\lambda y}\left(1-e^{-\lambda y}\right)^{\alpha-1} - \alpha T_0 e^{-\lambda T_0}\left(1-e^{-\lambda T_0}\right)^{\alpha-1}}{(1-e^{-\lambda y})^{\alpha}-(1-e^{-\lambda T_0})^{\alpha}} \\ &-(n-D-s)\frac{\alpha y e^{-\lambda y}\left(1-e^{-\lambda y}\right)^{\alpha-1}}{1-(1-e^{-\lambda y})^{\alpha}} = 0. \end{aligned} \tag{3.6}$$

From (3.4) and (3.6), we obtain the MLP of Y as

$$\widehat{Y}_{MLP} = -\frac{1}{\tilde{\lambda}}\ln\left[1-\left(\frac{(s-1)\tilde{\alpha}T_0 e^{-\tilde{\lambda}T_0}\left(1-e^{-\tilde{\lambda}T_0}\right)^{\tilde{\alpha}-1}}{\frac{D+1}{\tilde{\lambda}} - \sum_{i=1}^{D}\frac{x_i(1-\tilde{\alpha}e^{-\tilde{\lambda}x_i})}{1-e^{-\tilde{\lambda}x_i}}} + \left(1-e^{-\tilde{\lambda}T_0}\right)^{\tilde{\alpha}}\right)^{1/\tilde{\alpha}}\right] \tag{3.7}$$

where $(\tilde{\alpha},\tilde{\lambda})$ is PMLE of (α,λ) that can be obtained numerically from (3.5) and (3.6).

3.2 Conditional prediction approach

In conditional prediction approach, the conditional distribution of $Y = X_{s+D:n}$ given $\mathbf{X} = (X_1, ..., X_D)$ is applied to derive the predictors of Y. A statistic $\hat{Y}$ which is used to predict $Y = X_{s+D:n}$ is called a best unbiased predictor (BUP) of Y, if the predictor error $\hat{Y} - Y$ has a mean zero and its prediction error variance $\text{Var}(\hat{Y} - Y)$ is less than or equal to that of any other unbiased predictor of Y.

The BUP of Y is

$$\hat{Y}_{BUP} = E(Y|\mathbf{X}) = \int_{T_0}^{\infty} y f(y|\mathbf{x}, \alpha, \lambda) dy.$$

Using (2.5) and the binomial expansion

$$\left[\left(1 - e^{-\lambda y}\right)^{\alpha} - \left(1 - e^{-\lambda T_0}\right)^{\alpha}\right]^{s-1} =$$

$$\sum_{j=0}^{s-1} \binom{s-1}{j} (-1)^{s-j-1} \left[1 - \left(1 - e^{-\lambda T_0}\right)^{\alpha}\right]^{j} \left[1 - \left(1 - e^{-\lambda y}\right)^{\alpha}\right]^{s-j-1} \tag{3.8}$$

we obtain

$$\begin{aligned} \hat{Y}_{BUP} &\equiv I(T_0; \alpha, \lambda) \\ &= s\binom{n-D}{s} \sum_{j=0}^{s-1} \binom{s-1}{j} (-1)^{s-j-1} \left[1 - \left(1 - e^{-\lambda T_0}\right)^{\alpha}\right]^{-(n-D-j)} h_1(T_0), \end{aligned} \tag{3.9}$$

where

$$\begin{aligned} h_1(T_0) &= \int_{T_0}^{\infty} y\alpha\lambda e^{-\lambda y} \left(1 - e^{-\lambda y}\right)^{\alpha-1} \left[1 - \left(1 - e^{-\lambda y}\right)^{\alpha}\right]^{n-D-j-1} dy \\ &= -\frac{\alpha}{\lambda} \int_{0}^{e^{-\lambda T_0}} \ln(t)(1-t)^{\alpha-1}[1 - (1-t)^{\alpha}]^{n-D-j-1} dt. \end{aligned} \tag{3.10}$$

Because the parameters α and λ are unknown, they have to be estimated. Thus, one would replace them by their corresponding MLEs and obtain the BUP of Y.

Let us now consider conditional median predictor (CMP) which was first suggested by Raqab and Nagaraja (1995). A predictor $\hat{Y}$ is called the CMP of Y, if it is the median of the conditional distribution of Y given $\mathbf{X} = \mathbf{x}$, that is

$$P_{\theta}(Y \leq \hat{Y}|\mathbf{X} = \mathbf{x}) = P_{\theta}(Y \geq \hat{Y}|\mathbf{X} = \mathbf{x}). \tag{3.11}$$

Using the relation

$$P_{\alpha,\lambda}(Y \leq \hat{Y}|\mathbf{X}=\mathbf{x}) = P_{\alpha,\lambda}\left(1-\frac{1-\left(1-e^{-\lambda Y}\right)^{\alpha}}{1-\left(1-e^{-\lambda T_0}\right)^{\alpha}} \leq 1-\frac{1-\left(1-e^{-\lambda \hat{Y}}\right)^{\alpha}}{1-\left(1-e^{-\lambda T_0}\right)^{\alpha}}|\mathbf{X}=\mathbf{x}\right), \tag{3.12}$$

and using the fact that the distribution of $1-\frac{1-\left(1-e^{-\lambda Y}\right)^{\alpha}}{1-\left(1-e^{-\lambda T_0}\right)^{\alpha}}$ given $\mathbf{X}=\mathbf{x}$ is a $Beta(s, n-D-s+1)$ distribution, we obtain the CMP of Y as

$$\hat{Y}_{CMP} = -\frac{1}{\lambda}\ln\left[1-\left(1-[1-Med(B)]\left[1-\left(1-e^{-\lambda T_0}\right)^{\alpha}\right]\right)^{1/\alpha}\right], \tag{3.13}$$

where B has $Beta(s, n-D-s+1)$ distribution and $Med(B)$ stands for median of B. By substituting α and λ with their corresponding MLEs, we obtain the CMP of Y.

3.3 Classical prediction Intervals

In this section, we construct two prediction intervals (PI's) for $Y = X_s$ based on the Type I or Type-II censored sample $\mathbf{X} = (X_1, X_2, \cdots, X_D)$.

Let us define the random variable Z as

$$Z = 1-\frac{1-\left(1-e^{-\lambda Y}\right)^{\alpha}}{1-\left(1-e^{-\lambda T_0}\right)^{\alpha}}.$$

As mentioned before, the distribution of Z given $\mathbf{X}=\mathbf{x}$ is a $Beta(s, n-D-s+1)$ distribution. So, we can consider Z as a pivotal quantity to obtain the prediction interval for Y. Now, a $(1-\gamma)100\%$ PI for Y is $(L_1(\mathbf{X}), U_1(\mathbf{X}))$ where

$$L_1(\mathbf{X}) = -\frac{1}{\lambda}\ln\left[1-\left(1-\left[1-B_{\frac{\gamma}{2}}\right]\left[1-\left(1-e^{-\lambda T_0}\right)^{\alpha}\right]\right)^{1/\alpha}\right] \tag{3.14}$$

$$U_1(\mathbf{X}) = -\frac{1}{\lambda}\ln\left[1-\left(1-\left[1-B_{1-\frac{\gamma}{2}}\right]\left[1-\left(1-e^{-\lambda T_0}\right)^{\alpha}\right]\right)^{1/\alpha}\right] \tag{3.15}$$

where B_γ stands for 100γth percentile of $Beta(s, n-D-s+1)$ distribution. When α and λ are unknown, the parameters in (3.14) and (3.15), have to be estimated. For example, by replacing α and λ with their corresponding MLEs, the prediction limits for Y can be obtained.

Now let us consider another prediction interval for $Y = X_s$. The distribution of Z given $\mathbf{X}=\mathbf{x}$ is a $Beta(s, n-D-s+1)$ distribution with pdf

$$g(z) = \frac{z^{s-1}(1-z)^{n-D-s}}{Beta(s, n-D-s+1)}, \qquad 0 < z < 1,$$

which is a unimodal function of z, for $1 < s < n - D$. Therefore, the $(1-\gamma)100\%$ highest conditional density (HCD) prediction limits for Y are given by

$$L_2(\mathbf{X}) = -\frac{1}{\lambda}\ln\left[1 - \left(1 - [1 - w_1]\left[1 - \left(1 - e^{-\lambda T_0}\right)^{\alpha}\right]\right)^{1/\alpha}\right] \tag{3.16}$$

$$U_2(\mathbf{X}) = -\frac{1}{\lambda}\ln\left[1 - \left(1 - [1 - w_2]\left[1 - \left(1 - e^{-\lambda T_0}\right)^{\alpha}\right]\right)^{1/\alpha}\right] \tag{3.17}$$

where w_1 and w_2 are the simultaneous solutions of the following equations:

$$\int_{w_1}^{w_2} g(z)dz = 1 - \gamma \tag{3.18}$$

and

$$g(w_1) = g(w_2). \tag{3.19}$$

Now, we simplify Equations (3.18) and (3.19) as

$$B_{w_2}(s, n - D - s + 1) - B_{w_1}(s, n - D - s + 1) = 1 - \gamma, \tag{3.20}$$

and

$$\left(\frac{1 - w_2}{1 - w_1}\right)^{n-D-s} = \left(\frac{w_1}{w_2}\right)^{s-1}, \tag{3.21}$$

where

$$B_t(a, b) = \frac{1}{B(a, b)}\int_0^t x^{a-1}(1 - x)^{b-1}\, dx,$$

is the incomplete beta function. It is clear from (3.21), that the method cannot be used to construct the prediction interval when $s = 1$.

4 Bayesian prediction of future observations

In this section, we obtain the Bayesian predictors of $Y = X_{s+D},\ (s = 1, 2, ..., n - D)$ based on the observed hybrid censored sample $\mathbf{x} = (x_1, ..., x_D)$. For finding Bayesian predictors , it is assumed that α and λ each have independent $Gamma(a_1, b_1)$ and $Gamma(a_2, b_2)$ priors, respectively. Based on these priors, we can obtain the joint posterior density function of α and λ given the data as

$$\pi(\alpha, \lambda|\mathbf{x}) \propto g_1(\alpha|\lambda, \mathbf{x})g_2(\lambda|\mathbf{x}). \tag{4.1}$$

Here $g_2(\lambda|\mathbf{x})$ is a gamma density function with the shape and scale parameters as $D + a_2$ and $\sum_{i=1}^{D} x_i + b_2$, respectively. Further,

$$g_1(\alpha|\lambda, \mathbf{x}) \propto \frac{e^{-b_1\alpha}\alpha^{D+a_1-1}}{\left(\sum_{i=1}^{D} x_i + b_2\right)^{D+a_2}}\left[1 - \left(1 - e^{-\lambda T_0}\right)^{\alpha}\right]^{n-D}\prod_{i=1}^{D}\left(1 - e^{-\lambda x_i}\right)^{\alpha-1}, \tag{4.2}$$

which is a proper density function.

The Bayesian predictors are obtained from the predictive density function. The predictive density of $Y = X_{s+D}$ given $\mathbf{x}$ is

$$f_s^*(y|\mathbf{x}) = \int_0^\infty \int_0^\infty f(y|\mathbf{x}, \alpha, \lambda)\pi(\alpha, \lambda|\mathbf{x})d\alpha d\lambda \tag{4.3}$$

where $f(y|\mathbf{x}, \alpha, \lambda)$ is the conditional density of $Y = X_{s+D}$ given $\mathbf{x}$. By replacing (4.1) in (4.3), we obtain the predictive density function $f_s^*(y|\mathbf{x})$ as

$$f_s^*(y|\mathbf{x}) = \int_0^\infty \int_0^\infty f(y|\mathbf{x}, \alpha, \lambda)g_1(\alpha|\lambda, \mathbf{x})g_2(\lambda|\mathbf{x})d\alpha d\lambda. \tag{4.4}$$

Now, the Bayesian point (BP) predictor of Y under a squared error loss is

$$\widehat{Y}_{BP} = \int_{T_0}^\infty y f_s^*(y|\mathbf{x})dy. \tag{4.5}$$

It is not possible to compute (4.5) explicitly. Therefore, we propose here Gibbs sampling procedure to obtain the Bayes predictor. In the Gibbs sampling procedure, we can generate samples from the posterior density function $\pi(\alpha, \lambda|\mathbf{x})$ and in turn obtain the Bayes predictor and also the corresponding prediction intervals based on the generated posterior samples. For later use, we need the following result.

Theorem 1: The conditional distribution of α given the data, $g_1(\alpha|\lambda, \mathbf{x})$ is log-concave.

Proof: See the Appendix.

Now, using Theorem 1 and following the idea of Geman and Geman (1984), we propose the following algorithm to generate (α, λ) from the posterior density function.

1. Generate λ_1 from $g_2(.|\mathbf{x})$.
2. Generate α_1 from $g_1(.|\lambda_1, \mathbf{x})$ using the method proposed by Devroye (1984).
3. Repeat steps 1 and 2, M times and obtain $(\alpha_1, \lambda_1), \cdots, (\alpha_M, \lambda_M)$.

Now, using the generated samples $(\alpha_1, \lambda_1), \cdots, (\alpha_M, \lambda_M)$, the simulation consistent estimator of $f_s^*(y|\mathbf{x})$ is

$$\widehat{f_s^*}(y|\mathbf{x}) = \frac{1}{M}\sum_{i=1}^{M} f(y|\mathbf{x}, \alpha_i, \lambda_i). \tag{4.6}$$

Therefore, by replacing (4.6) in (4.5), the BP of Y can be approximated as

$$\hat{Y}_{BP} = \int_{T_0}^{\infty} y \left(\frac{1}{M} \sum_{i=1}^{M} f(y|\mathbf{x}, \alpha_i, \lambda_i) \right) dy \tag{4.7}$$

$$= \frac{1}{M} \sum_{i=1}^{M} \int_{T_0}^{\infty} y f(y|\mathbf{x}, \alpha_i, \lambda_i) dy \tag{4.8}$$

$$= \frac{1}{M} \sum_{i=1}^{M} I(T_0; \alpha_i, \lambda_i), \tag{4.9}$$

where $I(T_0, \alpha, \lambda)$ is defined in (3.9).

Bayesian prediction intervals are obtained from the Bayes predictive density $f^*(y|\mathbf{x})$. For some positive u, the survivor function of Y given $\mathbf{x}$ is

$$P(Y > u|\mathbf{x}) = \int_{u}^{\infty} f^*(y|\mathbf{x}) dy.$$

which by using (4.6) can be approximated as

$$P^(Y > u|\mathbf{x}) = \frac{1}{M} \sum_{i=1}^{M} \int_{u}^{\infty} f(y|\mathbf{x}, \alpha_i, \lambda_i) dy$$

$$= s\binom{n-D}{s} \frac{1}{M} \sum_{i=1}^{M} \sum_{j=0}^{s-1} \binom{s-1}{j} \frac{(-1)^{s-j-1}}{n-D-j} \left[\frac{1-\left(1-e^{-\lambda_i u}\right)^{\alpha_i}}{1-\left(1-e^{-\lambda_i T_0}\right)^{\alpha_i}} \right]^{n-D-j}.$$

Therefore, the $100(1-\gamma)\%$ Bayesian prediction interval for Y is $(L_3(\mathbf{X}), U_3(\mathbf{X}))$, where the prediction limits $L_3(\mathbf{X})$ and $U_3(\mathbf{X})$ can be obtained by solving the nonlinear equations

$$P(Y > L_3(\mathbf{x})|\mathbf{x}) = 1 - \frac{\gamma}{2}, \qquad P(Y > U_3(\mathbf{x})|\mathbf{x} = \frac{\gamma}{2}.$$

5 Real Data Analysis

In this section, we consider one real data set mainly to illustrate how the different methods perform when the underlying distribution is known to be GE. The following data below (Lawless, 1982, p. 267) are from an accelerator life test of 59 conductors. The data are

2.997	4.137	4.288	4.531	4.700	4.706	5.009	5.381	5.434	5.459
5.589	5.640	5.807	5.923	6.033	6.071	6.087	6.129	6.352	6.369
6.479	6.492	6.515	6.522	6.538	6.545	6.573	6.725	6.869	6.923
6.948	6.956	6.958	7.024	7.224	7.365	7.398	7.459	7.489	7.495
7.496	7.543	7.683	7.937	7.945	7.974	8.120	8.336	8.532	8.591
8.687	8.799	9.218	9.254	9.289	9.663	10.092	10.491	11.038	

A subset of the data set we will be using as samples (training data) and the rest will be used for verification purposes. Before progressing further, first we want to see whether GE model provides a good fit to the data set or not. We obtain the MLEs of the GE parameters as $\widehat{\alpha}$ =52.411 and $\widehat{\lambda}$ =0.642. It is observed that the Kolmogorov-Smirnov (K-S) distance and the corresponding p-value are, respectively,

$$\text{K-S} = 0.103, \qquad \text{and} \qquad p = 0.532.$$

Since p-value is quite high, we can say that the GE provides a good fit to the data set.

SCHEME 1: Let us consider the following sampling scheme: $r = 20$ and $T = 6$. Based on Type-I HCS, the sample is:

```
2.997  4.137  4.288  4.531  4.700  4.706  5.009  5.381  5.434  5.459
5.589  5.640  5.807  5.923.
```

In this case $D = 14$ and $T_0 = 6$. Now based on the above sample, we want to predict $X_{(s)}$, for different values of s, namely for $s = 1, 3, 5, 8$. The predicted values and the prediction intervals using different methods are computed and they are reported in Table 1. Note that for computing Bayesian predictions, since we do not have any prior information, we assumed that the priors on α and λ are improper, i.e. $a_1 = b_1 = a_2 = b_2 = 0$.

SCHEME 2: In this case we consider Type-II HCS with the same $r = 20$ and $T = 6$. The observed sample becomes

```
2.997  4.137  4.288  4.531  4.700  4.706  5.009  5.381  5.434  5.459
5.589  5.640  5.807  5.923  6.033  6.071  6.087  6.129  6.352
6.369.
```

Here $D = 20$ and $T_0 = 6.369$. Now based on the above Type-II HCS, the predicted value of $X_{(s)}$ and the prediction intervals are reported in Table 1.

Table 1. The values of point predictions and 95% PIs for $Y = X_s$.

Type-I HCS								
		Point Prediction				Interval Prediction		
s	Exact Value	MLP	BUP	CMP	BP	Pivotal Method	HCD Method	Bayesian Method
s=1	6.033	5.923	6.030	5.999	6.048	(5.926,6.307)	—	(5.602,6.352)
s=3	6.087	6.131	6.241	6.213	6.301	(5.993,6.651)	(5.958,6.578)	(5.461,6.725)
s=5	6.352	6.334	6.448	6.423	6.604	(6.106,6.937)	(6.067,6.875)	(5.988,7.205)
s=8	6.492	6.633	6.966	6.734	7.203	(6.313,7.331)	(6.275,7.280)	(6.163,7.464)
Type-II HCS								
s=1	6.479	6.369	6.464	6.436	6.521	(6.372,6.710)	—	(6.065,6.929)
s=3	6.515	6.557	6.653	6.626	6.852	(6.430,7.028)	(6.403,6.961)	(6.386,7.094)
s=5	6.538	6.745	6.843	6.819	7.162	(6.532,7.299)	(6.498,7.239)	(6.284,7.744)
s=8	6.725	7.030	7.156	7.112	7.802	(6.721,7.685)	(6.690,7.636)	(6.069,8.785)

From Table 1, it is observed that different point predictors are quite close to the true observations. Also, we observe that different prediction intervals are containing the true observations.

6 Numerical Comparisons

It is difficult to compare the performances of the different predictors and predictive intervals theoretically, as proposed in the previous sections. In this section, we present a simulation study to compare the performance of the different methods. The simulation is performed based on two types of hybrid censoring schemes, and for different T and different parameter values. We have kept $r = 10$ and $n = 20$ to be fixed throughout. The experiment can be briefly describe as follows. For given n and the parameter values, we generate a sample of size n from a GE(α, λ). Based on the sampling scheme we obtain the (training) sample from the whole sample as $\{x_1, \ldots, x_D\}$. Now based on the sample $\{x_1, \ldots, x_D\}$, we predict $Y = X_{s+D:n}$, and also obtain different predictive intervals. For computing Bayesian point and interval predictors, we assume two priors as follows:

$$\text{Prior 1:} \quad a_j = 0.0001, \quad b_j = 0.0001, \quad j = 1, 2,$$
$$\text{Prior 2:} \quad a_j = 1, \quad b_j = 3, \quad j = 1, 2.$$

The two priors have the same means. But, the variance of prior 2 is smaller than that of prior 1. So, the prior 2 is more informative than prior 1 We compare the performances of the MLP, BUP, CMP and Bayes predictors in terms of biases, mean squares prediction errors (MSPE) and Pitman closeness (PC) measure, based on 10000 replications. It may be mentioned the measure of Pitman closeness has received considerable attention in recent years, see for example Balakrishnan et al. (2011) and the references cited therein. The bias and MSPE of the different predictors are reported in Table 2, and the PC measures of the different predictors are reported in Table 4.

We have calculated the Bias, MSPE and PC as follows. Suppose $\hat{Y}_i$ is the prediction of $Y = X_{s+D:n}$ derived in the i-th iteration of simulation, where $i = 1, \ldots, N = 10000$, then the Bias and the MSPE computed as:

$$Bias = \frac{1}{N}\sum_{i=1}^{N}(\hat{Y}_i - Y) \tag{6.1}$$

and

$$MSPE = \frac{1}{N}\sum_{i=1}^{N}(\hat{Y}_i - Y)^2 \tag{6.2}$$

If we consider, $\hat{Y}_i^{(1)}$ and $\hat{Y}_i^{(2)}$ as the predictors of $Y = X_{s+D:n}$, obtained based on the first and second method, respectively, then PC for $\hat{Y}_i^{(1)}$ against $\hat{Y}_i^{(2)}$, can be calculated as follows:

$$PC = P(|\hat{Y}^{(1)} - Y| < |\hat{Y}^{(2)} - Y|) = \frac{1}{N}\sum_{i=1}^{N} I(|\hat{Y}_i^{(2)} - Y| - |\hat{Y}_i^{(1)} - Y|) \tag{6.3}$$

where $I(.)$ is the indicator function, and $I(x) = 1$ if $x > 0$ and 0, otherwise. We say that $\hat{Y}^{(1)}$ competes with $\hat{Y}^{(2)}$ if $PC > 0.5$.

We also compute three prediction intervals obtained using pivotal, HCD and Bayesian methods as discussed in Sections 3 and 4. We compared the performances of the three intervals in terms of their average lengths and coverage percentages based on 10000 replications. The results are reported in Table 3.

Table 2. The biases and MSPEs of point predictors for $n = 20$ and $r = 10$.

			$\alpha = 0.75$, $\lambda = 1$									
			$T = 0.40$					$T = 0.60$				
			BUP	MLP	CMP	BP		BUP	MLP	CMP	BP	
						Prior 1	Prior 2				Prior 1	Prior 2
Type-I HCS	s=1	Biase	0.016	0.021	0.018	0.024	0.020	0.019	0.024	0.021	0.034	0.030
		MSPE	0.003	0.004	0.004	0.006	0.004	0.002	0.002	0.002	0.006	0.003
	s=2	Biase	0.024	0.026	0.019	0.016	0.017	0.024	0.026	0.022	0.038	0.028
		MSPE	0.004	0.005	0.006	0.009	0.006	0.002	0.003	0.003	0.007	0.004
	s=3	Biase	0.026	0.030	0.023	0.026	0.024	0.031	0.029	0.028	0.028	0.026
		MSPE	0.006	0.007	0.009	0.010	0.008	0.004	0.006	0.006	0.009	0.006
	s=4	Biase	0.028	0.027	0.031	0.041	0.028	0.042	0.024	0.029	0.029	0.027
		MSPE	0.008	0.008	0.010	0.012	0.010	0.005	0.006	0.008	0.011	0.009
	s=5	Biase	0.034	0.029	0.025	0.027	0.019	0.038	0.034	0.026	0.019	0.017
		MSPE	0.010	0.012	0.013	0.015	0.013	0.007	0.009	0.010	0.014	0.010
Type-II HCS	s=1	Biase	0.016	0.026	0.027	0.029	0.024	0.015	0.020	0.021	0.013	0.016
		MSPE	0.001	0.002	0.003	0.005	0.002	0.001	0.001	0.001	0.003	0.001
	s=2	Biase	0.026	0.027	0.021	0.016	0.019	0.019	0.015	0.024	0.013	0.019
		MSPE	0.002	0.004	0.005	0.008	0.005	0.001	0.002	0.002	0.006	0.004
	s=3	Biase	0.027	0.029	0.026	0.042	0.046	0.016	0.024	0.019	0.024	0.035
		MSPE	0.004	0.006	0.007	0.009	0.007	0.002	0.004	0.005	0.007	0.006
	s=4	Biase	0.019	0.015	0.022	0.031	0.037	0.019	0.027	0.030	0.064	0.050
		MSPE	0.005	0.007	0.008	0.011	0.010	0.003	0.005	0.006	0.010	0.007
	s=5	Biase	0.029	0.025	0.020	0.030	0.026	0.020	0.024	0.021	0.024	0.014
		MSPE	0.008	0.010	0.012	0.014	0.012	0.005	0.007	0.009	0.012	0.009
			$\alpha = 1.5$, $\lambda = 1$									
			$T = 0.80$					$T = 1.20$				
			BUP	MLP	CMP	BP		BUP	MLP	CMP	BP	
						Prior 1	Prior 2				Prior 1	Prior 2
Type-I HCS	s=1	Biase	0.030	0.036	0.039	0.049	0.037	0.035	0.029	0.041	0.024	0.028
		MSPE	0.005	0.007	0.007	0.012	0.009	0.004	0.006	0.007	0.010	0.008
	s=2	Biase	0.033	0.028	0.046	0.037	0.029	0.024	0.028	0.037	0.028	0.027
		MSPE	0.008	0.010	0.011	0.016	0.012	0.006	0.008	0.008	0.014	0.011
	s=3	Biase	0.045	0.056	0.028	0.049	0.050	0.064	0.052	0.034	0.042	0.046
		MSPE	0.011	0.014	0.016	0.019	0.013	0.009	0.010	0.012	0.021	0.018
	s=4	Biase	0.059	0.067	0.065	0.034	0.022	0.029	0.034	0.035	0.075	0.046
		MSPE	0.013	0.016	0.019	0.025	0.021	0.010	0.013	0.016	0.021	0.018
	s=5	Biase	0.062	0.067	0.046	0.024	0.029	0.049	0.057	0.061	0.075	0.060
		MSPE	0.016	0.018	0.022	0.028	0.026	0.013	0.019	0.021	0.027	0.023
Type-II HCS	s=1	Biase	0.058	0.064	0.046	0.072	0.068	0.042	0.038	0.048	0.067	0.062
		MSPE	0.005	0.006	0.007	0.010	0.008	0.003	0.004	0.006	0.009	0.008
	s=2	Biase	0.041	0.049	0.046	0.034	0.037	0.045	0.036	0.047	0.045	0.054
		MSPE	0.007	0.008	0.008	0.011	0.010	0.005	0.007	0.008	0.010	0.009
	s=3	Biase	0.033	0.037	0.035	0.028	0.026	0.056	0.034	0.045	0.060	0.054
		MSPE	0.009	0.010	0.013	0.017	0.015	0.008	0.009	0.011	0.016	0.013
	s=4	Biase	0.051	0.052	0.046	0.057	0.053	0.078	0.082	0.064	0.067	0.072
		MSPE	0.011	0.013	0.014	0.020	0.016	0.009	0.012	0.015	0.018	0.014
	s=5	Biase	0.067	0.066	0.055	0.058	0.056	0.034	0.038	0.064	0.052	0.042
		MSPE	0.012	0.018	0.020	0.025	0.022	0.022	0.016	0.018	0.024	0.021

Table 3. Average lengths and coverage probabilities of 95% PIs for $n = 20$ and $r = 10$.

			$\alpha = 0.75$, $\lambda = 1$							
			$T = 0.4$				$T = 0.6$			
			Pivotal Method	HCD Method	Bayesian Prior 1	Bayesian Prior 2	Pivotal Method	HCD Method	Bayesian Prior 1	Bayesian Prior 2
Type-I HCS	s=1	Length	0.125	—	0.112	0.103	0.116	—	0.107	0.091
		CP	0.934	—	0.935	0.938	0.933	—	0.935	0.937
	s=2	Length	0.203	0.180	0.195	0.188	0.161	0.134	0.153	0.146
		CP	0.936	0.941	0.938	0.940	0.935	0.940	0.936	0.937
	s=3	Length	0.247	0.219	0.235	0.230	0.199	0.164	0.189	0.173
		CP	0.938	0.941	0.938	0.941	0.937	0.941	0.936	0.938
	s=4	Length	0.298	0.269	0.286	0.276	0.234	0.206	0.222	0.216
		CP	0.942	0.945	0.942	0.944	0.940	0.944	0.940	0.941
	s=5	Length	0.352	0.329	0.348	0.337	0.279	0.242	0.266	0.254
		CP	0.943	0.947	0.945	0.945	0.942	0.945	0.943	0.944
Type-II HCS	s=1	Length	0.133	—	0.124	0.116	0.108	—	0.096	0.088
		CP	0.933	—	0.936	0.934	0.929	—	0.934	0.937
	s=2	Length	0.196	0.169	0.181	0.175	0.148	0.119	0.138	0.130
		CP	0.935	0.938	0.937	0.934	0.930	0.934	0.933	0.931
	s=3	Length	0.264	0.231	0.254	0.242	0.194	0.160	0.175	0.168
		CP	0.938	0.941	0.938	0.940	0.935	0.938	0.936	0.936
	s=4	Length	0.306	0.281	0.297	0.291	0.254	0.224	0.244	0.238
		CP	0.939	0.943	0.943	0.943	0.938	0.943	0.940	0.941
	s=5	Length	0.336	0.310	0.328	0.321	0.266	0.238	0.255	0.247
		CP	0.941	0.946	0.942	0.945	0.941	0.944	0.941	0.943
			$\alpha = 1.5$, $\lambda = 1$							
			$T = 0.8$				$T = 1.2$			
Type-I HCS	s=1	Length	0.167	—	0.154	0.147	0.142	—	0.132	0.125
		CP	0.937	—	0.939	0.940	0.935	—	0.936	0.938
	s=2	Length	0.266	0.241	0.258	0.252	0.206	0.174	0.194	0.187
		CP	0.938	0.942	0.939	0.941	0.936	0.940	0.938	0.939
	s=3	Length	0.322	0.301	0.320	0.311	0.261	0.238	0.250	0.244
		CP	0.938	0.943	0.940	0.941	0.938	0.942	0.939	0.941
	s=4	Length	0.405	0.371	0.394	0.382	0.317	0.284	0.310	0.301
		CP	0.944	0.949	0.947	0.949	0.944	0.947	0.945	0.945
	s=5	Length	0.470	0.451	0.464	0.460	0.344	0.318	0.329	0.324
		CP	0.947	0.955	0.951	0.954	0.945	0.952	0.950	0.950
Type-II HCS	s=1	Length	0.161	—	0.150	0.142	0.138	—	0.126	0.120
		CP	0.935	—	0.938	0.937	0.933	—	0.934	0.936
	s=2	Length	0.231	0.212	0.224	0.219	0.226	0.201	0.219	0.211
		CP	0.938	0.942	0.939	0.941	0.935	0.938	0.935	0.937
	s=3	Length	0.298	0.268	0.284	0.278	0.287	0.261	0.280	0.274
		CP	0.941	0.945	0.943	0.945	0.937	0.941	0.938	0.938
	s=4	Length	0.394	0.367	0.388	0.379	0.311	0.294	0.304	0.300
		CP	0.942	0.946	0.945	0.945	0.940	0.945	0.942	0.944
	s=5	Length	0.461	0.449	0.459	0.455	0.339	0.312	0.330	0.328
		CP	0.944	0.949	0.946	0.948	0.942	0.947	0.945	0.946

Table 4. PC comparison of different predictors for $n = 20$ and $r = 10$.

		$\alpha = 0.75$, $\lambda = 1$							
		$T = 0.40$				$T = 0.60$			
		BUP vs MLP	BUP vs CMP	MLP vs CMP	BP(Prior 2) vs BP(Prior 1)	BUP vs MLP	BUP vs CMP	MLP vs CMP	BP(Prior 2) vs BP(Prior 1)
Type-I HCS	s=1	0.545	0.568	0.526	0.649	0.584	0.604	0.529	0.684
	s=2	0.612	0.649	0.586	0.682	0.648	0.682	0.629	0.719
	s=3	0.671	0.697	0.641	0.729	0.706	0.729	0.687	0.780
	s=4	0.712	0.748	0.730	0.772	0.738	0.759	0.720	0.816
	s=5	0.746	0.759	0.741	0.791	0.784	0.798	0.758	0.842
Type-II HCS	s=1	0.527	0.558	0.521	0.613	0.542	0.579	0.534	0.640
	s=2	0.549	0.584	0.531	0.663	0.576	0.597	0.564	0.687
	s=3	0.604	0.627	0.589	0.705	0.637	0.684	0.610	0.729
	s=4	0.642	0.681	0.624	0.734	0.685	0.705	0.659	0.779
	s=5	0.693	0.713	0.673	0.751	0.725	0.746	0.706	0.792
		$\alpha = 1.5$, $\lambda = 1$							
		$T = 0.80$				$T = 1.20$			
Type-I HCS	s=1	0.546	0.568	0.537	0.629	0.564	0.579	0.543	0.613
	s=2	0.576	0.592	0.550	0.671	0.591	0.628	0.571	0.649
	s=3	0.608	0.631	0.582	0.703	0.638	0.662	0.616	0.682
	s=4	0.628	0.654	0.603	0.741	0.652	0.689	0.662	0.726
	s=5	0.659	0.682	0.631	0.768	0.694	0.726	0.672	0.752
Type-II HCS	s=1	0.528	0.549	0.520	0.608	0.534	0.550	0.528	0.597
	s=2	0.549	0.573	0534.	0.614	0.552	0.576	0.534	0.621
	s=3	0.568	0.593	0.551	0.629	0.581	0.608	0.571	0.657
	s=4	0.586	0.616	0.572	0.684	0.620	0.657	0.604	0.703
	s=5	0.634	0.652	0.612	0.711	0.659	0.689	0.639	0.726

Some of the points are quite clear from this simulation experiments.

- From Table 2, it is observed the MSPEs increase as s increases for all the prediction methods.
- For the same α, λ and T, the MSPEs of the different predictors for Type-II HCS are smaller than the corresponding Type-I HCS.
- In terms of MSPEs, BUP is slightly better than the MLP and CMP.
- In terms of MSPEs, the Bayes predictors under prior 2 work better than the Bayes predictor under prior 1.
- From Table 4, it is observed that based on PC measure BUP is better than the MLP and CMP. Also based on PC measure, the BP predictor under prior 2 works better than the BP predictor under prior 1.
- From Table 3, it is observed that the coverage probabilities for all prediction intervals are very close to the nominal level.

- For the same α, λ and T, the average lengths and coverage percentages increase as s increases for all the methods.
- From Table 3, the HCD method works better than the the pivotal and Bayesian methods in terms of both average lengths and coverage probabilities. Also, the Bayesian prediction intervals under Prior 2 are shorter than the Bayesian prediction intervals under Prior 1.

Therefore, based on the simulation results here, we propose the following: for prediction of future observation BUP may be used, and for prediction interval, for $s > 1$, we propose to use the HCD method, and for $s = 1$, Bayesian method may be used.

7 Conclusions

In this paper we have considered the prediction of future observations and the associated prediction interval based on Type-I HCS or Type-II HCS samples, when the lifetime distributions of the experimental units follow GE distribution. We have used three different classical predictors namely BUP, MLP and CMP. We have also considered Bayesian predictor under squared error loss function. We then compared these different point predictors using biases, MSPE and PC. Comparing all the points, we recommend to use BUP for predicting the future observation. To calculate the prediction intervals we have used three methods namely HCD, Pivotal and Bayesian methods. It is observed HCD method cannot be used when $s = 1$. Three proposed methods maintain the coverage percentages. HCD method performs slightly better than the Pivotal method and Bayesian method for $s > 1$.

Acknowledgements:

The authors would like to thank the referees and the associate editor for their very constructive suggestions.

Appendix

The conditional density of α given λ and the data is

$$g_1(\alpha|\lambda, \mathbf{x}) \propto \frac{e^{-b_1\alpha}\alpha^{D+a_1-1}}{\left(\sum_{i=1}^{D} x_i + b_2\right)^{D+a_2}} \left[1 - \left(1 - e^{-\lambda T_0}\right)^{\alpha}\right]^{n-D} \prod_{i=1}^{D} \left(1 - e^{-\lambda x_i}\right)^{\alpha-1}.$$

The log-likelihood of $g_1(\alpha|\lambda, \mathbf{x})$ is

$$\begin{aligned} \ln g_1(\alpha|\lambda, \mathbf{x}) &\propto -b_1\alpha + (D + a_1 - 1)\ln(\alpha) + (\alpha - 1)\sum_{i=1}^{D} \ln[1 - e^{-\lambda x_i}] \\ &+ (n - D)\ln\left[1 - (1 - e^{-\lambda T_0})^{\alpha}\right] \end{aligned} \quad (7.1)$$

Using (7.1), we have

$$\frac{d^2 \ln g_1(\alpha|\lambda, \mathbf{x})}{d\alpha^2} = -\frac{D + a_1 - 1}{\alpha^2} - (n - D)\frac{\ln^2\left[1 - e^{-\lambda T_0}\right](1 - e^{-\lambda T_0})^{\alpha}}{\left(1 - (1 - e^{-\lambda T_0})^{\alpha}\right)^2} \leq 0.$$

Therefore the result follows.

References

[1] A. Asgharzadeh, R. Valiollahi, D. Kundu. (2013). Prediction for future failures in Weibull distribution under hybrid censoring. *Journal of Statistical Computation and Simulation*, http://dx.doi.org/10.1080/00949655.2013.848451.

[2] N. Balakrishnan, K.F. Davies, K.F., J.P. Keating, R.L. Mason (2011). "Pitman closeness, monotonicity and consistency of best linear unbiased and in variant estimators for exponential distribution under Type-II censoring", *Journal of Statistical Computation and Simulation*, vol. 81, 985 - 999.

[3] N. Balakrishnan, E. Beutner, E. Cramer (2010), "Exact two-sample non- parametric confidence, prediction, and tolerance intervals based on ordinary and progressively type-II right censored data", *Test*, vol. 19, 68 - 91.

[4] I. Basak, P. Basak and N. Balakrishnan (2006), "On some predictors of times to failures of censored items in progressively censored sample", *Computational Statistics and Data Analysis*, vol. 50, 1313 - 1337

[5] N. Balakrishnan and D. Kundu (2012). "Hybrid censoring: models, inferential results and applications" (with discussion), *Computational Statistics and Data Analysis*, vol. 57, pp. 166-209.

[6] S. Chen and G. K. Bhattacharya, (1988). "Exact confidence bounds for an exponential parameter under hybrid censoring", *Communications in Statistics- Theory and Methods*, vol. 17, pp. 1857-1870.

[7] A. Childs, B. Chandrasekhar, N. Balakrishnan, and D. Kundu, (2003). "Exact likelihood inference based on type-I and type-II hybrid censored samples from the exponential distribution", *Annals of the Institute of Statistical Mathematics*, vol. 55, pp. 319-330.

[8] P. Dellaportas and D. Wright (1991). "Numerical prediction for the two-parameter Generalized exponential distribution". *The Statistician*, vol. 40, pp. 365 - 372.

[9] N. Draper and I. Guttman, (1987). "Bayesian analysis of hybrid life tests with exponential failure times", *Annals of the Institute of Statistical Mathematics*, vol. 39, pp. 219-225.

[10] N. Ebrahimi, (1986). "Estimating the parameter of an exponential distribution from hybrid life test", *Journal of Statistical Planning and Inference*, vol. 14, pp. 255-261.

[11] N. Ebrahimi, (1992). "Prediction intervals for future failures in exponential distribution under hybrid censoring", *IEEE Transactions on Reliability*, vol. 41, pp. 127-132.

[12] B. Epstein, (1954). "life tests in the exponential case", *Annals of Mathematical Statistics*, vol. 25, pp. 555 - 564.

[13] K. Fairbanks, R. Madison, and R. Dykstra, (1982). "A confidence interval for an exponential parameter from a hybrid life test", *Journal of the American Statistical Association*, vol. 77, pp. 137-140.

[14] S. Geman, and A. Geman (1984). Stochastic relaxation, Gibbs distributions and the Bayesian restoration of images. *IEEE Trans Pattern Anal Mach Intell*, vol. 6, pp. 721-740.

[15] H. S. Jeong, J. I. Park, and B. J. Yum, (1996). "Development of (r, T) hybrid sampling plans for exponential lifetime distributions", *Journal of Applied Statistics*, vol. 23, pp. 601-607.

[16] K. S. Kaminsky and L. S. Rhodin, (1985)." Maximum likelihood prediction". *Annals of the Institute of Statistical Mathematics*, vol. 37, 507-517

[17] D. Kundu, (2007). "On hybrid censoring Generalized exponential distribution", *Journal of Statistical Planning and Inferencee*, vol. 137, pp. 2127-2142.

[18] D. Kundu and A. Banerjee, (2008). "Inference Based on Type-II Hybrid Censored Data From a Generalized exponential Distribution," *IEEE Transaction on Reliability*, vol. 57, no. 2, pp. 369-378.

[19] D. Kundu and R. D. Gupta, (1988). "Hybrid censoring schemes with exponential failure distribution", *Communications in Statistics- Theory and Methods*, vol. 27, pp. 3065-3083.

[20] D. Kundu and M. Z. Raqab (2012). "Bayesian inference and prediction of order statistics for a Type-II censored Weibull distribution", *Journal of Statistical Planning and Inference*, vol. 142, pp. 41-47.

[21] J. F. Lawless. (1982). "*Statistical Models and Methods for Life Time Data*", John Wiley, NewYork.

[22] M. Z. Raqab and H. N. Nagaraja, (1995). "On some predictors of future order statistics," *Metron*, vol. 53(12), pp. 185-204.

[23] M. Z. Raqab, A. Asgharzadeh, and R. Valiollahi, (2010). "Prediction for Pareto distribution based on progressively Type-II censored samples". *Computational Statistics and Data Analysis*, vol. 54, 1732-1743

[24] R. L. Smith (1997), "Statistics for exceptional athletics record: letter to the editor". *Applied Statistics*, vol. 46, pp. 123-127.

[25] R. L. Smith (1999), "Bayesian and frequentist approaches to parametric predictive inference". *Bayesian Statistics* Vol. 6, Editors: Bernado, J. M., Berger, J.O., Dawid, A. P., Smith, A. F. M., 589 - 612, (with discussion), Oxford University Press, Oxford.