Generalized Bayes Prediction Study Based on Joint Type-II Censoring

Yahia Abdel-Aty 1,2, Mohamed Kayid 3 and Ghadah Alomani 4,*

1 Department of Mathematics, College of Science, Taibah University, Al-Madinah Al-Munawarah 30002, Saudi Arabia; ymahmoudawad@taibahu.edu.sa
2 Department of Mathematics, Faculty of Science, Al-Azhar University, Nasr City 11884, Egypt
3 Department of Statistics and Operations Research, College of Science, King Saud University, P.O. Box 2455, Riyadh 11451, Saudi Arabia; drkayid@ksu.edu.sa
4 Department of Mathematical Sciences, College of Science, Princess Nourah bint Abdulrahman University, P.O. Box 84428, Riyadh 11671, Saudi Arabia
* Correspondence: gaalomani@pnu.edu.sa

Abstract: In this paper, the problem of predicting future failure times based on a jointly type-II censored sample from k exponential populations is considered. The Bayesian prediction intervals and point predictors were then obtained. Generalized Bayes is a Bayesian study based on a learning rate parameter. This study investigated the effects of the learning rate parameters on the prediction results. The loss functions of squared error, Linex, and general entropy were used as point predictors. Monte Carlo simulations were performed to show the effectiveness of the learning rate parameter in improving the results of prediction intervals and point predictors.

Keywords: learning rate parameter; exponential distribution; joint type-II censoring; point predictor; prediction intervals; squared-error loss; Linex loss; general entropy loss

MSC: 62F10; 62F15

1. Introduction and Motivations

Generalized Bayes is a Bayesian study based on a learning rate parameter (η > 0) as a power of the likelihood function \( L(\theta; data) \). The traditional Bayes framework is obtained for η = 1, and we demonstrate the effect of the learning rate parameter on the prediction results. That is, if the prior distribution of the parameter \( \theta \) is \( \pi(\theta) \) then the generalized Bayes posterior distribution for \( \theta \) is

\[
\pi^*(\theta \mid data) \propto L^\eta(\theta; data) \pi(\theta), \quad \theta \in \Theta, \quad \eta > 0.
\] (1)

For more details on the generalized Bayes method and the choice of the value of the rate parameter, we refer the reader to [1–11]. In a special way, the choice of the learning rate \( \eta \) was studied in [3–6] by a so-called safe Bayes algorithm based on the minimization of a sequential risk measure. Another learning rate selection method considers the two different information-matching strategies proposed in [7,8]. In addition, a generalized Bayes estimation based on a joint censored sample of type II from k exponential populations using different values of the learning rate parameter was studied in [11]. An exact inference method based on maximum likelihood estimates (MLEs) was developed in [12], and its performance was compared with that of approximate, Bayesian, and bootstrap methods. The joint progressive censoring type II and the expected number of failures for two populations under the joint progressive censoring type II were introduced and studied by [13]. In contrast, the exact likelihood inference for two exponential populations under joint progressive censoring of type II was studied in [14], and some precise results were obtained based on the maximum likelihood estimates developed by...
[15]. Exact likelihood inference for two populations of two-parameter exponential distributions under type II joint censoring was studied by [16].

One might be interested in predicting future failures using a joint type II censored sample. To accomplish this, prediction points or intervals should be determined. Bayesian prediction bounds for future observations based on certain distributions were discussed by several authors. A study of Bayesian estimation and prediction based on a joint censored sample of type II from two exponential populations was presented by [17]. Prediction (various classical and Bayesian point predictors) for future failures in the Weibull distribution under hybrid censoring was studied by [18]. A Bayesian prediction based on generalized order statistics with multiple censoring of type II was developed by [19].

The main objective of this study is to predict future failures based on a joint type-II censoring scheme for \(k\)-exponential populations when censoring is performed on \(k\) samples in a combined manner. Suppose that products from \(k\) different lines are produced in the same factory, and \(k\) independent samples of size \(n_h, 1 \leq h \leq k\) are selected from these \(k\) lines and simultaneously placed in a lifetime experiment. To reduce the cost and time of the experiment, the experimenter may decide to stop the lifetime test when a certain number \((r)\) of failures occurs. The nature of the problem and the distributions used in our study are presented below.

Suppose \(\{X^n_h, h = 1, ..., k\}\) are \(k\)-samples, where \(X^n_h = \{X_{h1}, X_{h2}, ..., X_{hn_h}\}\) are the lifetimes of \(n_h\) samples of product line \(A_h\) and are assumed to be independent and identically distributed (iid) random variables from a population with a probability density function (pdf) \(f_h(x)\) and a cumulative distribution function (cdf) \(F_h(x)\).

Furthermore, let \(N = \sum_{i=1}^k n_i\) be the total sample size, and let \(r\) be the total number of observed failures. Let \(W_1 \leq \cdots \leq W_k\) denote the order statistics of \(N\) random variables \(\{X^n_h, h = 1, ..., k\}\). Under the joint type-II censoring scheme for the \(k\)-samples, the observable data then consist of \((\delta, W)\), where \(W = (W_1, ..., W_r), W_i \in \{X^n_{h_i}, h_i = 1, ..., k\}\), with \(r < N\) being a pre-fixed integer and \(\delta = (\delta_1(h), ..., \delta_r(h))\) associated to \((h_1, ..., h_r)\) is defined by

\[
\delta_i(h) = \begin{cases} 1, & \text{if } h = h_i \\ 0, & \text{otherwise.} \end{cases}
\] (2)

Letting \(M_r(h) = \sum_{i=1}^r \delta_i(h)\) denote the number of \(X_h\)-failures in \(W\) and \(r = \sum_{i=1}^k M_r(h)\), the joint density function of \((\delta, W)\) is given by

\[
f(\delta, W) = \prod_{i=1}^k c_r(\bar{F}_h(W_r))^{n_h - M_r(h)} \prod_{i=1}^r \prod_{h=1}^k (f_h(W_i))^{\delta_i(h)}
\] (3)

where \(\bar{F}_h = 1 - F_h\) is the survival function of \(h^{th}\) population and \(c_r = \frac{n_h!}{(n_h - M_r(h))!}\).

For any continuous variables \(Y_1 \leq \cdots \leq Y_n\), the joint density function of \(Y_{r+s}, Y_{r+s}, r < s \leq n\) is given by

\[
f(y_1, ..., y_r, y_s) = \frac{n!}{(s - r - 1)! (n - s)!} \prod_{i=1}^r f(y_i) \prod_{j=r+1}^s f(y_j).
\]

Here, \((W_1, ..., W_r, W_s) \equiv (W, W_s)\) linked with the discrete variables \((\delta_1, ..., \delta_r, \delta_s) \equiv (\delta, \delta_s)\).

Then the joint density function of \((\delta, \delta_s, W, W_s), r < s \leq N\) is given by

\[
f(\delta, \delta_s, W, W_s) = \sum_{Q_{r+s}} \prod_{h=1}^k c_1(h) (\bar{F}_h(W_r) - \bar{F}_h(W_s))^{M_{r+s}(h)}
\]

\[
\times (\bar{F}_h(W_s))^{n_h s} (f_h(W_s))^{\delta_s(h)} \prod_{i=1}^r \prod_{h=1}^k (f_h(W_i))^{\delta_i(h)}
\] (4)

where
\[ w_r < w_s \leq w_N, \ M_{rs}(h) = M_{s-1}(h) - M_r(h), \ \bar{n}_{hs} = n_h - M_s(h), \ c_{1h} = \frac{n_{h!}}{M_{rs}(h) \bar{n}_{hs}!}. \ 
\]

\[ \sum_{\delta_{r+1} = 0}^{1} \cdots \sum_{\delta_s = 0}^{1} \text{with} \ Q_s = \{ \delta(h) = (\delta_{r+1}, \ldots, \delta_s), \text{for } 1 \leq h \leq k \}. \]

The conditional density function of \( W_s \) given \((\delta, W) = (\delta, w)\), is given by

\[
f(w_s | \delta, w) = \sum_{Q_s} \prod_{h=1}^{k} c_h(f_h(w_s))^{\delta_s(h)} \left\{ \bar{F}_h(w_r) - \bar{F}_h(w_s) \right\}^{M_{rs}(h)} \frac{\bar{n}_{hs}!}{(\bar{F}_h(w_r))^{\bar{n}_{hs}+l_h} (\bar{F}_h(w_s))^{\bar{n}_{hs}+l_h}}
\]

where

\[ c_h = \frac{\bar{n}_{hr}!}{M_{rs}(h) \bar{n}_{hs}!} \ a_{lh} = (-1)^{l_h} \left( \frac{M_{rs}(h)}{l_h} \right). \]

In addition, when the \( k \) populations are exponential, the pdf is given by

\[
f_h(w) = \theta_h \exp(-\theta_h w), \text{and cdf } F_h(w) = 1 - \exp(-\theta_h w),
\]

where \( w > 0, \theta_h > 0; \ 1 \leq h \leq k \).

Then, the likelihood function in (3) becomes

\[
f(\Theta, \delta, w) = \prod_{h=1}^{k} c_r(\exp(-\theta_h w_r))^{\bar{n}_{hr}} \prod_{l=1}^{r} \prod_{h=1}^{k} \left\{ \theta_h \exp(-\theta_h w_l) \right\}^{\delta_l(h)}
\]

\[ = \prod_{h=1}^{k} c_r \theta_h^{M_r(h)} \exp\{-\theta_h u_h\}
\]

where \( \Theta = (\theta_1, \ldots, \theta_k) \) and \( u_h = \sum_{i=1}^{h} w_i \delta_i(h) + w_r \bar{n}_{hr} \).

Substituting (6) into (5), we obtain the conditional density function of \( W_s \), given \((\delta, w) = (\delta, w)\),

\[
f(w_s | \delta, w) = \sum_{Q_s} \prod_{h=1}^{k} \theta_h^{\delta_s(h)} \sum_{l=0}^{M_{rs}(h)} C_h \exp\{-\theta_h D_h(w_s - w_r)\}
\]

where \( C_h = c_h a_{lh} \) \( D_h = \bar{n}_{h(s-1)} + l_h \) and \( \bar{n}_{hs} + \delta_s(h) = \bar{n}_{h(s-1)} \).

Some special cases of the conditional density function are described as follows:

**Case 1:**

Suppose that \( k - 1 \) is the number of samples satisfy \( M_r(h) = n_h \), but only one sample from the \( k \) samples say \( \{X_{q1}^q\} \) satisfies \( M_r(q) < n_q \) or \( \{w_1, \ldots, w_k\} \in \{X_{q1}^q\} \).

Under Case 1, the conditional density function of \( W_s \) given \((\delta, w) = (\delta, w)\) becomes

\[
f_1(w_s | \delta, w) = \frac{\bar{n}_{qr}!}{(\bar{n}_{qr} - s + r)!} \sum_{i=0}^{s-r-1} a_{iq} \theta_q \exp\{-\theta_q D_q(w_s - w_r)\}
\]

where \( a_{iq} = (-1)^q \binom{s-r-1}{l_q} \) \( D_q = \bar{n}_{qr} - s + r + l_q + 1 \), \( w_r < w_s \leq x_{qnq} \).

**Case 2:**

Suppose that \( k - R \) is the number of samples satisfy \( M_r(h_i) = n_{h_i} \) for \( h_i = 1, \ldots, k \); and \( R < k \) is the number of samples satisfy \( M_r(h_j) < n_{h_j} \) for \( h_j = 1, \ldots, k; h_i \neq h_j \), equivalently, \( W_r > \max \{X_{q1}^q, h_i = 1, \ldots, k\} \) but \( W_r \in \{X_{q1}^q, h_j = 1, \ldots, k; h_i \neq h_j\} \).

Under Case 2, let us just consider the \( R \) samples, where \( q = 1, \ldots, R \), then the conditional density function of \( W_s \) given \((\delta, w) = (\delta, w)\) becomes
\[ f_2(w_s | \delta, w) = \sum_{Q_s'} \prod_{q=1}^{R} \delta_s^{(q)} \sum_{l_q=0}^{M_{rs}(q)} C_q \exp\{-\theta_q D_q (w_s - w_r)\} \]

where \( Q_s' = (\delta_1(q), \ldots, \delta_s(q)) \) for \( 1 \leq q \leq R \), \( C_q = c_q^s a_{l_q} \), \( D_q = \bar{n}_{q(s-1)} + l_q \) and \( \bar{n}_{qs} + \delta_s(q) = \bar{n}_{q(s-1)} \).

The remainder of this article is organized as follows: Section 2 presents the generalized Bayesian and Bayesian prediction points and intervals using squared error, Linex, and general entropy loss functions in the point predictor. A numerical study of the results from Section 2 is presented in Section 3. Finally, we conclude the paper in Section 4.

2. Generalized Bayes Prediction

In this section, we introduce the concept of generalized Bayesian prediction, which is an investigation of Bayesian prediction under the influence of a learning rate parameter \( \eta > 0 \). To apply the concept of generalized Bayesian prediction to a prediction study, we give a brief description of generalized Bayesian prediction based on a learning rate parameter \( \eta > 0 \). A scheme for predicting a sample based on joint censoring of type II samples from \( k \) exponential distributions is presented. The main goal is to obtain the point predictors and prediction intervals given at the end of this section.

2.1. Generalized Bayes

The parameters \( \Theta \) are assumed to be unknown, we may consider the conjugate prior distributions of \( \Theta \) as independent gamma prior distributions, i.e., \( \theta_h \sim \text{Gam}(a_h, b_h) \). Hence, the joint prior distribution of \( \Theta \) is given by

\[ \pi(\Theta) = \prod_{h=1}^{k} \pi_h(\theta_h), \]  

where

\[ \pi_h(\theta_h) = \frac{b_h^{a_h}}{\Gamma(a_h)} \theta_h^{a_h-1} \exp\{-b_h \theta_h\}, \]  

and \( \Gamma(\cdot) \) denotes the complete gamma function.

Combining (7) and (10) after raising (7) to the power \( \eta \), the posterior joint density function of \( \Theta \) is then

\[ \pi^*(\Theta | \delta, w) = \prod_{h=1}^{k} \frac{\left(u_h \eta + b_h\right)^{\eta M_r(h) + a_h} \theta_h^{\eta M_r(h) + a_h - 1}}{\Gamma(\eta M_r(h) + a_h)} \exp\{-\theta_h (u_h \eta + b_h)\}, \]

\[ = \prod_{h=1}^{k} \frac{\xi_h^{\mu_h}}{\Gamma(\mu_h)} \exp\{-\theta_h \xi_h\}, \]

where \( \xi_h = u_h \eta + b_h, \mu_h = \eta M_r(h) + a_h. \)

Since \( \pi_h \) is a conjugate prior, where \( \theta_h \sim \text{Gam}(a_h, b_h) \), then it follows that the posterior density function of \( \Theta_h | \delta, w \) is \( \text{Gam}(\eta M_r(h) + a_h, u_h \eta + b_h) \).

2.2. One Sample Prediction

A sample prediction scheme for the case of the joint censoring of samples from two exponential distributions was studied in [17], and then three cases for the future failures were derived; where in the first case, the future predicted failure surly belongs to \( X^1_n \) failures if \( M_r(1) < n_1, M_r(2) = n_2, \) in the second case, the future predicted failure surly belongs to \( X^2_n \) failures if \( M_r(2) < n_2, M_r(1) = n_1, \) and in the third case, it is unknown to which sample the future predicted failure belongs. Here, we generalize the results reported in [17] and examine two special cases in addition to the general case.
In the general case, the size of any sample is greater than the number of observed failures; that is, \( M_r(h) < n_h \equiv w_r < x_{h+1} \) for \( h = 1, \ldots, k \). The first special case arises when all future values (predictors) belong to only one sample and the observations of the remaining \( k - 1 \) samples are less than \( w_r \). The second special case arises when all future values (predictors) belong to some samples and all observations of the other samples are less than \( w_r \). The forms of all functions related to the second special case are similar to those related to the general case; therefore, we will introduce only the general case and the first special case.

For the general case, to predict \( w_r \) for \( r < s \leq N \) based on the observed data \((\delta, w)\), we use the conditional density function (9). Let us define the following integral:

\[
I^s_h(\delta(h)) = \frac{\xi h}{\Gamma(\mu h)} \int_0^\infty \theta h^{\mu h - 1 + \delta_h} \exp\{-\theta h[\xi h + D h(w_s - w_r)]\} d\theta h
\]

(12)

where \( \delta h(\delta) = 1 \) for \( \delta h(\delta) = 0 \).

Since \( \frac{a^b}{r(b)} \int_0^\infty x^b \exp(-x(a+c)) dx = \frac{a^b}{r(b)} \left( \frac{r(b+1)}{a+c} \right) = \frac{b}{a} \left( 1 + \frac{c}{a} \right)^{(b+1)}, \)

\[
\frac{a^b}{r(b)} \int_0^\infty x^{b-1} \exp(-x(a+c)) dx = \frac{a^b}{r(b)} \left( \frac{r(b)}{a+c} \right) = \left( 1 + \frac{c}{a} \right)^{-b}.
\]

Using (8), (12), and (13), the Bayesian predictive density function of \( W_s \) given \((\delta, w) = (\delta, w)\) becomes

\[
f_B(w_s | \delta, w) = \int_0^\infty f(w_s | \delta, w) \pi^*(\Theta | \delta, w) d\Theta
\]

\[
= \sum_{q_s} \prod_{h=1}^k \left\{ \sum_{l_h=0}^{M_{rs}} \frac{c_h \mu h}{\Gamma(\mu h)} \int_0^\infty \theta h^{\mu h - 1 + \delta_h} \exp\{-\theta h[\xi h + D h(w_s - w_r)]\} d\theta h \right\}
\]

(13)

where

\[
\prod_{h=1}^k \int_0^\infty f(\Theta h) d\Theta h = \int_0^\infty \cdots \int_0^\infty f(\Theta 1) \cdots f(\Theta k) d\Theta 1 \cdots d\Theta k.
\]

Under Case 1, the Bayesian predictive density function of \( W_s \) given \((\delta, w) = (\delta, w)\) becomes

\[
f_{1B}(w_s | \delta, w) = \frac{\bar{q}_q r \mu_q}{(\bar{q}_q r - s + l)} \sum_{l_q=0}^{s-r-1} a_{l_q} \left( 1 + \frac{D_{q}(w_s - w_r)}{\xi q} \right)^{-(\mu q + 1)},
\]

(14)

where \( \xi q = u_q \eta + b_q, \mu_q = \eta M_r(q) + a_q, w_r < w_s \leq x_{qn-q} \).

2.3. Bayesian Point Predictors

For the point predictor, we considered three types of loss functions:

(i). The squared error loss function (SE), which is classified as a symmetric function, is given by

\[
L_{SE}(\varphi^*, \varphi) \propto (\varphi^* - \varphi)^2,
\]

where \( \varphi^* \) is an estimate of \( \varphi \).

(ii). The Linex loss function, which is asymmetric, is given by
(iii). The generalization of the entropy (GE) loss function is

\[ L_{GE}(\varphi^*, \varphi) \propto \left( \frac{\varphi^*}{\varphi} \right)^c - c \ln \left( \frac{\varphi^*}{\varphi} \right) - 1, \quad c \neq 0. \]

It is worth noting that the Bayes estimates under the GE loss function coincide with those under the SE loss function when \( c = -1 \). However, when \( c = 1, -2 \), the Bayes estimates under GE become those under the weighted squared error loss function and the precautionary loss function, respectively.

Now, the Bayesian point predictors \( W_r, r < s \leq N \), under different loss functions (SE, Linex, and GE) can be obtained using the predictive density function (14), which are denoted, respectively, by \( W_{SP}, W_{LP}, W_{EP} \) and given as follows:

\[
W_{SP} = E(W_s | \delta, w) = \int_0^\infty w_s f_B(w_s | \delta, w) dw_s
\]

\[
= \sum_{k=1}^{\infty} \prod_{h=1}^{k} \prod_{l_h=0}^{M_{Rh}} \frac{C_h^k \left( \sum_{v=1}^{k} \int_0^\infty w_s (l_v^1 \prod_{q=1, q \neq v}^{k} l_q^0) dw_s \right)}{c}
\]

\[
W_{LP} = \frac{1}{r} \ln \left[ \int_0^\infty e^{-rW_s} f_B(w_s | \delta, w) dw_s \right]
\]

\[
= \frac{1}{r} \ln \left[ \sum_{k=1}^{\infty} \prod_{h=1}^{k} \prod_{l_h=0}^{M_{Rh}} \frac{C_h^k \left( \sum_{v=1}^{k} \int_0^\infty e^{-rW_s} (l_v^1 \prod_{q=1, q \neq v}^{k} l_q^0) dw_s \right)}{c} \right]
\]

\[
W_{EP} = \left[ E(W_s^{-c} | \delta, w) \right]^{\frac{-1}{c}}
\]

Under Case 1, \( W_{SP}, W_{LP}, \) and \( W_{EP} \) are, respectively, given by

\[
W_{SP} = \frac{\bar{r}q_{r}^1 \mu_q}{(\bar{r}q_{r} - s + r)} \sum_{l_q=0}^{s-r-1} a_{l_q} \int_0^\infty w_s (1 + \frac{D_q(w_s - w_r)}{\xi_q})^{-(\mu_q + 1)} dw_s
\]

\[
W_{LP} = \frac{1}{r} \ln \left[ \frac{\bar{r}q_{r}^1 \mu_q}{(\bar{r}q_{r} - s + r)} \sum_{l_q=0}^{s-r-1} a_{l_q} \int_0^\infty e^{-rW_s} (1 + \frac{D_q(w_s - w_r)}{\xi_q})^{-(\mu_q + 1)} dw_s \right]
\]

\[
W_{EP} = \left[ \frac{\bar{r}q_{r}^1 \mu_q}{(\bar{r}q_{r} - s + r)} \sum_{l_q=0}^{s-r-1} a_{l_q} \int_0^\infty w_s^{-c} (1 + \frac{D_q(w_s - w_r)}{\xi_q})^{-(\mu_q + 1)} dw_s \right]^{\frac{-1}{c}}.
\]

The above equations are solved numerically to obtain the predictors \( W_{SP}, W_{LP}, \) and \( W_{EP} \).

2.4. Prediction Interval
The predictive survival function of $W_s$ is given by:

$$
\bar{F}_B(t) = P(W_s > t \mid \delta, w) = \int_t^{\infty} f_B(w_s \mid \delta, w) \, dw_s
$$

$$
= \sum_{q_{s-1}} \prod_{h=1}^k \sum_{l_h=0}^{M_{rs}} C_h \left[ \sum_{q=1}^k \int_t^{\infty} \left( \prod_{q_{s-1}} \right) \prod_{q=1}^{r} l_q^0 \right] \, dw_s.
$$

(22)

Numerical integration is required to obtain the predictive survival function in Equation (23).

In Case 1, the predictive survival function of $W_s$ is given by

$$
\bar{F}_{1B}(t) = P(W_s > t \mid \delta, w) = \int_t^{\infty} f_{1B}(w_s \mid \delta, w) \, dw_s
$$

$$
= \frac{n_{qr}}{(n_{rq}-s+r)!} \sum_{r=0}^{s-r-1} \sum_{q=1}^k a_q \left( 1 + \frac{D_q(t-w_r)}{\zeta_q} \right)^{-\mu_q}.
$$

(20)

The Bayesian predictive bounds of a two-sided equi-tailed 100(1 - γ)% interval for $W_s, r < s \leq N$, can be obtained by solving the following two equations numerically,

$$
\bar{F}(L \mid \delta, w) = 1 - \frac{\gamma}{2}, \quad \bar{F}(U \mid \delta, w) = \frac{\gamma}{2}.
$$

(21)

3. Numerical Study

In this section, the results of the Monte Carlo simulation study are conducted to evaluate the performance of the prediction study derived in the previous section, and an example is presented to illustrate the prediction methods discussed here.

3.1. Simulation Study

We considered three samples from three populations with $(n_1, n_2, n_3, r)$ for choices (10, 10, 10, 25) and (15, 15, 15, 40). In Case 1, we choose the exponential parameters $(\theta_1, \theta_2, \theta_3)$ as (2, 1, 0.1) based on the hyperparameters represented by $\Delta = (a_1, b_1, a_2, b_2, a_3, b_3)$, where $\Delta = \Delta_1 = (2, 1, 2, 1, 10)$.

In the general case, we choose the exponential parameters $(\theta_1, \theta_2, \theta_3)$ as (2, 2.5, 3) based on the hyperparameters $\Delta = \Delta_2 = (2, 1.5, 2, 3, 1)$.

For the generalized Bayesian study, three values are chosen for the learning rate parameter $\eta = 1, 2, 5$, and 10,000 repetitions are used for the Monte Carlo simulations.

The mean observations values of the three generated samples $X_1, X_2,$ and $X_3$, and their joint sample $W$, using 10,000 repetitions, are presented in Tables 1–4, where the underlined values are greater than $w_r$.

**Table 1. The mean observations values for $(n_1, n_2, n_3, r) = (10, 10, 10, 25), \Delta = \Delta_1$ (Case 1).**

<table>
<thead>
<tr>
<th>Sample</th>
<th>Data</th>
</tr>
</thead>
<tbody>
<tr>
<td>$X_1$</td>
<td>0.0498, 0.1058, 0.1674, 0.2391, 0.3225, 0.4226, 0.5477, 0.7159, 0.9664, 1.4685.</td>
</tr>
<tr>
<td>$X_2$</td>
<td>0.1005, 0.2130, 0.3363, 0.4780, 0.6444, 0.8434, 1.0905, 1.4248, 1.9253, 2.9195.</td>
</tr>
</tbody>
</table>

**Ordered data $(w, h)$, $r = 25$.**

(0.0498, 1), (0.1005, 2), (0.1058, 1), (0.1674, 1), (0.2391, 2), (0.3225, 1), (0.3225, 1), (0.3363, 2), (0.4226, 1), (0.4780, 2), (0.5477, 1), (0.6444, 2), (0.7159, 1), (0.8434, 2), (0.9664, 1), (1.01284, 3), (1.0905, 2), (1.4248, 2), (1.4685, 1), (1.9253, 2), (2.1473, 3), (2.9195, 2), (3.3872, 3), (4.8117, 3), (6.4598, 3).

**Table 2. The mean observations values for $(n_1, n_2, n_3, r) = (15, 15, 15, 40), \Delta = \Delta_1$ (Case 1).**
Sample | Data
---|---
$X_1$ | 0.0333, 0.0693, 0.1078, 0.1487, 0.1944, 0.2449, 0.3004, 0.3646, 0.4363, 0.5206, 0.6208, 0.7479, 0.9146, 1.1641, 1.6557.
$X_2$ | 0.0663, 0.1388, 0.2164, 0.2991, 0.3916, 0.4914, 0.6029, 0.7289, 0.8690, 1.0349, 1.2367, 1.4884, 1.8181, 2.316, 3.3304.
$X_3$ | 0.6610, 1.3818, 2.1570, 2.9962, 3.9136, 4.9102, 6.0317, 7.2834, 8.7249, 10.3833, 12.3990, 14.9019, 18.2287, 23.2551, 33.3370.

Ordered data $(w, h_i)$, $r = 40$

<table>
<thead>
<tr>
<th>Sample</th>
<th>Data</th>
</tr>
</thead>
<tbody>
<tr>
<td>$X_1$</td>
<td>(0.0333, 1), (0.0663, 2), (0.0693, 1), (0.1078, 1), (0.1388, 2), (0.1487, 1), (0.1944, 1), (0.2164, 2), (0.2449, 1), (0.2991, 2), (0.3004, 1), (0.3646, 1), (0.3916, 2), (0.4363, 1), (0.4914, 2), (0.5206, 1), (0.6029, 2), (0.6208, 1), (0.6610, 3), (0.7289, 2), (0.7479, 1), (0.8690, 2), (0.9146, 1), (1.0349, 2), (1.1641, 1), (1.2367, 2), (1.3818, 3), (1.4884, 2), (1.6557, 1), (1.8181, 2), (2.1570, 3), (2.316, 2), (2.9962, 3), (3.304, 2), (3.9136, 3), (4.9102, 3), (6.0317, 3), (7.2834, 3), (8.7249, 3), (10.3833, 3).</td>
</tr>
</tbody>
</table>

We notice from Tables 1 and 2 that, the future values come only from sample $X_3$.

Table 3. The mean observations values for $(n_1, n_2, n_3, r) = (10, 10, 10, 25)$, $\Delta = \Delta_2$, (general case).

Sample | Data
---|---
$X_1$ | 0.0508, 0.1062, 0.1693, 0.2417, 0.3241, 0.4219, 0.5451, 0.7110, 0.9583, 1.4663.
$X_2$ | 0.0401, 0.0844, 0.1338, 0.1907, 0.2575, 0.3384, 0.4395, 0.5727, 0.7732, 1.1655.
$X_3$ | 0.0334, 0.0703, 0.1127, 0.1604, 0.2167, 0.2839, 0.3671, 0.4778, 0.6474, 0.9823.

Ordered data $(w, h_i)$, $r = 25$

<table>
<thead>
<tr>
<th>Sample</th>
<th>Data</th>
</tr>
</thead>
<tbody>
<tr>
<td>$X_1$</td>
<td>(0.0334, 3), (0.0401, 2), (0.0508, 1), (0.0703, 3), (0.0844, 2), (0.1062, 1), (0.1127, 3), (0.1338, 2), (0.1604, 3), (0.1693, 1), (0.1907, 2), (0.2167, 3), (0.2417, 1), (0.2575, 2), (0.2839, 3), (0.3241, 1), (0.3384, 2), (0.3671, 3), (0.4219, 1), (0.4395, 2), (0.4778, 3), (0.5451, 1), (0.5727, 2), (0.6474, 3), (0.7110, 3).</td>
</tr>
</tbody>
</table>

Table 4. The mean observations values for $(n_1, n_2, n_3, r) = (15, 15, 15, 40)$, $\Delta = \Delta_2$, (general case).

Sample | Data
---|---
$X_1$ | 0.0326, 0.0681, 0.1068, 0.1487, 0.1944, 0.2449, 0.3003, 0.3632, 0.4344, 0.5186, 0.6195, 0.7433, 0.9125, 1.1605, 1.6591.
$X_2$ | 0.0269, 0.0556, 0.0864, 0.1193, 0.1557, 0.1965, 0.2414, 0.2913, 0.3484, 0.4137, 0.4943, 0.5941, 0.7264, 0.9289, 1.3301.
$X_3$ | 0.0220, 0.0459, 0.0713, 0.0988, 0.1295, 0.1624, 0.1989, 0.2399, 0.2872, 0.3424, 0.4084, 0.4915, 0.6032, 0.7706, 1.1045.

Ordered data $(w, h_i)$, $r = 40$

<table>
<thead>
<tr>
<th>Sample</th>
<th>Data</th>
</tr>
</thead>
<tbody>
<tr>
<td>$X_1$</td>
<td>(0.0220, 3), (0.0269, 2), (0.0326, 1), (0.0459, 3), (0.0556, 2), (0.0681, 1), (0.0713, 3), (0.0864, 2), (0.0988, 3), (0.1068, 1), (0.1193, 2), (0.1295, 1), (0.1487, 1), (0.1557, 2), (0.1624, 3), (0.1944, 1), (0.1965, 2), (0.1989, 3), (0.2399, 3), (0.2414, 2), (0.2449, 1), (0.2872, 3), (0.2913, 2), (0.3003, 1), (0.3424, 3), (0.3484, 2), (0.3632, 1), (0.4084, 3), (0.4137, 2), (0.4344, 1), (0.4915, 3), (0.4943, 2), (0.5186, 1), (0.5941, 2), (0.6032, 3), (0.6195, 1), (0.7264, 2), (0.7433, 1), (0.7706, 3), (0.9125, 1).</td>
</tr>
</tbody>
</table>

We notice from Tables 3 and 4 that the future values come from the three samples.

For $(n_1, n_2, n_3, r) = (10, 10, 10, 25)$, under Case 1, we use (19), (20), and (21) to calculate the mean squared prediction errors (MSPEs) of the point predictors $(W_{sp}, W_{lp}, \text{and } W_{rp})$ for $s = 26, \ldots, 30$, where $r = 0.1, 0.5$; $c = 0.1, 0.5$, and the results are presented in Table 5.

Table 5. MSPE of point predictions for $\eta = 1, 2, 5; \Delta = \Delta_1$ in Case 1.
The results of the MSPEs in the general case are calculated using (16), (17), and (18), and shown in Table 6.

Table 6. MSPEs of point predictions for η = 1, 2, 5; Δ = Δ2.

<table>
<thead>
<tr>
<th>(n₁,n₂,n₃,r)</th>
<th>s</th>
<th>SP</th>
<th>LP</th>
<th>EP</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td></td>
<td>τ = 0.1</td>
<td>τ = 0.5</td>
<td>c = 0.1</td>
</tr>
<tr>
<td>(10,10,10,25)</td>
<td>26</td>
<td>0.0238</td>
<td>0.0215</td>
<td>0.0211</td>
</tr>
<tr>
<td></td>
<td>27</td>
<td>0.2623</td>
<td>0.2425</td>
<td>0.2383</td>
</tr>
<tr>
<td></td>
<td>28</td>
<td>0.8924</td>
<td>0.8914</td>
<td>0.8767</td>
</tr>
<tr>
<td></td>
<td>29</td>
<td>1.4967</td>
<td>1.4687</td>
<td>1.3263</td>
</tr>
<tr>
<td></td>
<td>30</td>
<td>2.1610</td>
<td>2.0341</td>
<td>1.9864</td>
</tr>
</tbody>
</table>

For (n₁,n₂,n₃,r) = (10,10,10,25) and (n₁,n₂,n₃,r) = (15,15,15,40), the results of the prediction bounds of Wₛ, s = 26, ..., 30 and s = 41, ..., 45, respectively, are calculated using (23) and (24) in Case 1, then are presented in Table 7.
Table 7. Lower and upper 95% prediction bounds for $W_\alpha$ in Case 1, for different choices of $n_1, n_2, n_3, r$ and $\Delta = \Delta_1$.

<table>
<thead>
<tr>
<th>$(n_1, n_2, n_3, r)$</th>
<th>$s$</th>
<th>$\eta = 1$</th>
<th>$\eta = 2$</th>
<th>$\eta = 5$</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>$L$</td>
<td>$U$</td>
<td>$L$</td>
<td>$U$</td>
</tr>
<tr>
<td>(10, 10, 10, 25)</td>
<td>26</td>
<td>7.1426</td>
<td>9.3742</td>
<td>7.1517</td>
</tr>
<tr>
<td></td>
<td>27</td>
<td>9.8765</td>
<td>13.6496</td>
<td>9.9164</td>
</tr>
<tr>
<td></td>
<td>28</td>
<td>12.6289</td>
<td>21.3289</td>
<td>12.6714</td>
</tr>
<tr>
<td></td>
<td>29</td>
<td>16.7643</td>
<td>29.1496</td>
<td>16.8653</td>
</tr>
<tr>
<td></td>
<td>30</td>
<td>25.7658</td>
<td>49.5654</td>
<td>25.8585</td>
</tr>
<tr>
<td>(15, 15, 15, 40)</td>
<td>41</td>
<td>11.6539</td>
<td>13.8764</td>
<td>11.9152</td>
</tr>
<tr>
<td></td>
<td>42</td>
<td>12.9876</td>
<td>17.6824</td>
<td>12.9256</td>
</tr>
<tr>
<td></td>
<td>43</td>
<td>15.2879</td>
<td>23.1859</td>
<td>15.5429</td>
</tr>
<tr>
<td></td>
<td>44</td>
<td>20.3289</td>
<td>33.9126</td>
<td>20.9721</td>
</tr>
<tr>
<td></td>
<td>45</td>
<td>29.1289</td>
<td>51.1610</td>
<td>29.5289</td>
</tr>
</tbody>
</table>

Table 8 presents the prediction bounds using (22), and (24) to show the results of the general case.

Table 8. Lower and upper 95% prediction bounds for $W_\alpha$, for different choices of $n_1, n_2, n_3, r$ and $\Delta = \Delta_2$.

<table>
<thead>
<tr>
<th>$(n_1, n_2, n_3, r)$</th>
<th>$s$</th>
<th>$\eta = 1$</th>
<th>$\eta = 2$</th>
<th>$\eta = 5$</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>$L$</td>
<td>$U$</td>
<td>$L$</td>
<td>$U$</td>
</tr>
<tr>
<td>(10, 10, 10, 25)</td>
<td>26</td>
<td>0.6528</td>
<td>0.8889</td>
<td>0.6814</td>
</tr>
<tr>
<td></td>
<td>27</td>
<td>0.7589</td>
<td>1.1096</td>
<td>0.7768</td>
</tr>
<tr>
<td></td>
<td>28</td>
<td>0.8934</td>
<td>1.9583</td>
<td>0.9214</td>
</tr>
<tr>
<td></td>
<td>29</td>
<td>0.9789</td>
<td>3.5696</td>
<td>0.9901</td>
</tr>
<tr>
<td></td>
<td>30</td>
<td>1.1289</td>
<td>5.5610</td>
<td>1.2541</td>
</tr>
<tr>
<td>(15, 15, 15, 40)</td>
<td>41</td>
<td>0.7902</td>
<td>1.8846</td>
<td>0.7968</td>
</tr>
<tr>
<td></td>
<td>42</td>
<td>0.8674</td>
<td>2.2354</td>
<td>0.8776</td>
</tr>
<tr>
<td></td>
<td>43</td>
<td>0.9282</td>
<td>3.7583</td>
<td>0.9245</td>
</tr>
<tr>
<td></td>
<td>44</td>
<td>0.9949</td>
<td>4.5610</td>
<td>1.0237</td>
</tr>
<tr>
<td></td>
<td>45</td>
<td>1.2367</td>
<td>5.9810</td>
<td>1.2568</td>
</tr>
</tbody>
</table>

3.2. Illustrative Example

To illustrate the usefulness of the results developed in the previous sections, we consider three samples of size $n_1 = n_2 = n_3 = 10$ from Nelson’s data (groups 1, 4, and 5) corresponding to the breakdown of an insulating fluid subjected to a high-stress load (see [20], p. 462). These breakdown times, referred to here as samples $X_i$, $i = 1, 2, 3$, are jointly type-II censored data in the form of $(w, h_i)$ obtained from these three samples with $r = 24$ and are shown in Table 9.

Table 9. The failure time data for $X_1, X_2, \text{ and } X_3$, and their order $(w, h_i)$, where $\delta_{h_i} = 1$.

<table>
<thead>
<tr>
<th>Sample</th>
<th>Data</th>
</tr>
</thead>
<tbody>
<tr>
<td>$X_1$</td>
<td>1.89, 4.03, 1.54, 0.31, 0.66, 1.7, 2.17, 1.82, 9.99, 2.24</td>
</tr>
<tr>
<td>$X_2$</td>
<td>1.17, 3.87, 2.8, 0.7, 3.82, 0.02, 0.5, 3.72, 0.06, 3.57</td>
</tr>
<tr>
<td>$X_3$</td>
<td>8.11, 3.17, 5.55, 0.80, 0.20, 1.13, 6.63, 1.08, 2.44, 0.78</td>
</tr>
</tbody>
</table>

Ordered data $(w, h_i)$

$(0.02, 2), (0.06, 2), (0.20, 3), (0.31, 1), (0.50, 2), (0.66, 1), (0.70, 2), (0.78, 3), (0.80, 3), (1.08, 3)$
Using (16), (17), (18), (22), and (24), the MSPEs of the point predictors and prediction intervals of \( w, s = 25, \ldots, 30 \) are calculated and presented in Table 10 using \( \eta = 1, 2, 5 \) and \( \Delta = \Delta_3 = (1,2,6,1,2,1,3) \); and \( \tau = 0.1, 0.5 \) and \( c = 0.1, 0.5 \).

<table>
<thead>
<tr>
<th>( s )</th>
<th>( (w, h) )</th>
<th>( W_{SP} )</th>
<th>( W_{LP} )</th>
<th>( W_{LP} )</th>
<th>( (L, U) )</th>
</tr>
</thead>
<tbody>
<tr>
<td>Exact value</td>
<td>( \tau = 0.1 )</td>
<td>( \tau = 0.5 )</td>
<td>( c = 0.1 )</td>
<td>( c = 0.5 )</td>
<td></td>
</tr>
<tr>
<td>25</td>
<td>(3.87,2)</td>
<td>5.678</td>
<td>3.7287</td>
<td>3.7287</td>
<td>3.7789</td>
</tr>
<tr>
<td>26</td>
<td>(4.03,1)</td>
<td>6.4653</td>
<td>4.5364</td>
<td>4.4653</td>
<td>4.6532</td>
</tr>
<tr>
<td>27</td>
<td>(5.55,3)</td>
<td>5.8972</td>
<td>5.4676</td>
<td>5.3864</td>
<td>5.4943</td>
</tr>
</tbody>
</table>

Table 10. The values of point predictors and 95% prediction interval of \( W \), for \( \eta = 1, 2, 5 \); \( \Delta = \Delta_3 \).

### 4. Conclusions

In this study, we examined the effects of learning rate parameters on prediction results. We used Monte Carlo simulations to show the effectiveness of the learning rate parameter in improving the results of prediction intervals and point predictors. Formally, we considered a joint type-II censoring scheme in which the lifetimes of the three populations have exponential distributions. We determined the MSPEs of the point predictors and prediction intervals using different values for the learning rate parameter \( \eta \) and different values for the parameters of the losses in both the simulation study and the illustrative example. From all tables in this prediction study, it can be seen that the results improve with increasing the loss parameters \( c, \tau \) and learning rate parameter \( \eta \). In the simulation study, a comparison of the results in Tables 5 and 6 shows that the results in Table 5 are better, and the length of the prediction intervals in Table 8 is smaller than those in Table 7 because the observed values used in Table 7 are larger than those used in Table 8. The results of the illustrative example improve with larger values of loss parameters and learning rate parameter. So we conclude that the results of the prediction study became better as learning rate parameter increased. However, in both studies, the lengths of the prediction intervals increased for the larger future lifetimes. It may be interesting to examine this work using a different type of censoring.
Author Contributions: Conceptualization, Y.A.-A.; methodology, Y.A.-A. and M.K.; software, G.A.; validation, G.A.; formal analysis, M.K.; resources, Y.A.-A.; writing—original draft, Y.A.-A. and M.K.; writing—review & editing, G.A.; supervision, M.K.; project administration, Y.A.-A. and G.A. All authors have read and agreed to the published version of the manuscript.

Funding: Princess Nourah bint Abdulrahman University Researchers Supporting Project number (PNURSP2023R226), Princess Nourah bint Abdulrahman University, Riyadh, Saudi Arabia.

Data Availability Statement: The data used to support the findings of this study are included in the article.

Acknowledgments: The authors thank the three anonymous reviewers and the editor for their constructive criticism and valuable suggestions, which have greatly improved the presentation and explanations in this article. The authors extend their sincere appreciation to Princess Nourah bint Abdulrahman University Researchers Supporting Project Number (PNURSP2023R226), Princess Nourah bint Abdulrahman University, Riyadh, Saudi Arabia.

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

References

Disclaimer/Publisher’s Note: The statements, opinions and data contained in all publications are solely those of the individual author(s) and contributor(s) and not of MDPI and/or the editor(s). MDPI and/or the editor(s) disclaim responsibility for any injury to people or property resulting from any ideas, methods, instructions or products referred to in the content.