Statistics and Probability Letters 81 (2011) 360–370

Contents lists available at ScienceDirect

Statistics and Probability Letters

journal homepage: www.elsevier.com/locate/stapro

Point prediction of future order statistics from an exponential distribution

S.M.T.K. MirMostafaee*, Jafar Ahmadi

Department of Statistics, Ordered and Spatial Data Center of Excellence, Ferdowsi University of Mashhad, P. O. Box 1159, Mashhad, 91775, Iran

A R T I C L E   I N F O

Article history:
Received 31 July 2010
Received in revised form 22 November 2010
Accepted 22 November 2010
Available online 30 November 2010

MSC:
primary 62G30
secondary 62E15

Keywords:
BLIP
BLUP
MLP
MSPE
Record values

A B S T R A C T

Two-sample point prediction is considered for a two-parameter exponential distribution. Several point predictors such as the best unbiased predictor, best invariant predictor and maximum likelihood predictor are obtained for future order statistics on the basis of observed record values in two cases: where the location parameter is known and unknown. These predictors are compared in the sense of their mean squared prediction errors. Finally, some numerical results are given to illustrate the proposed procedures.

1. Introduction and preliminaries

Let $X_1, X_2, \ldots$ be an infinite sequence of random variables and $Z_n = \max\{X_1, \ldots, X_n\}$ for $n \geq 1$; then an observation X_j is called an upper record value if $Z_j > Z_{j-1},\ j > 1$. For a detailed review on applications of record values, see Arnold et al. (1998). One application of record values arises in association with a minimal repair process. Such a process can be explained briefly as follows. Consider the sequence of failure times of a technical system with one repairable component. After each failure, this component will immediately be repaired and restored to its previous condition. The times of repairs and replacements are very low and therefore neglected. If F denotes the cumulative distribution function (cdf) of the lifetime of that component, then the minimal repair times possess the same joint distribution as record values based on F (for more details see for example Shaked and Shanthikumar, 1994). Record values also have the same distribution as epoch times of some non-homogeneous Poisson process (NHPP; see e.g. Pellerey et al., 2000 for more details).

Next, let $Y_1, \ldots, Y_m$ be a sample of size m, independent of the X_n-sequence, and $Y_{j:m}$, $1 \leq j \leq m$, be the corresponding jth order statistic. Order statistics arise in many practical situations as well as the reliability of systems. It is well-known that a system is called a k-out-of-m system if it consists of m components functioning satisfactorily provided that at least k $(\leq m)$ components function. If the lifetimes of the components are independently distributed, then the lifetime of the system coincides with that of the $(m-k+1)$th order statistic from the underlying distribution. Therefore, order statistics play a key role in studying the lifetimes of such systems. See Arnold et al. (1992) and David and Nagaraja (2003) for more details concerning the applications of order statistics.

* Corresponding author.
E-mail addresses: ta_mi182@stu-um.ac.ir (S.M.T.K. MirMostafaee), ahmadi-j@um.ac.ir (J. Ahmadi).

0167-7152/$ – see front matter
doi:10.1016/j.spl.2010.11.017

Several researchers have studied prediction problems concerning order statistics and record values; e.g. Dunsmore (1983) investigated how to predict future record statistics, Hsieh (1997) studied the prediction of future order statistics from a Weibull distribution, and Raqab and Balakrishnan (2008) discussed the problem of prediction of future record values in nonparametric settings. Recently, Ahmadi and Balakrishnan (2010) considered the problem of predicting order statistics (record values) on the basis of record values (order statistics) and obtained several nonparametric prediction intervals. Ahmadi and MirMostafaee (2009) focused on the same problem and derived interval predictors for order statistics on the basis of record values from a two-parameter exponential distribution. Moreover, Ahmadi et al. (2010) studied the nonparametric outer and inner prediction intervals for future record intervals based on observed order statistics. In this paper, the main purpose is to find point predictors for future order statistics based on observed censored record values when the underlying distribution is a two-parameter exponential distribution.

Throughout this paper, we denote by $\mathrm{Exp}(\mu, \sigma)$ a two-parameter exponential distribution with probability density function (pdf)

$$f(x; \mu, \sigma) = \frac{1}{\sigma} \mathrm{e}^{-\frac{x-\mu}{\sigma}}, \quad x \geq \mu,\ \sigma > 0,\ -\infty < \mu < \infty. \tag{1}$$

Let $R_1, \ldots, R_n$ be the first n record values from $\mathrm{Exp}(\mu, \sigma)$ but where we have failed to observe the first $(s-1)$ records, namely we have observed $\mathbf{R}_{s,n} = (R_s, R_{s+1}, \ldots, R_n)$, $s - 1 < n$. Let $Y_{j:m}$ be the jth order statistic in a future sample of size m from the same distribution. In this paper we wish, on the basis of the observed $\mathbf{R}_{s,n} = (R_s, R_{s+1}, \ldots, R_n)$, to find point predictors for $Y_{j:m}$. In the following sections we obtain best unbiased predictors (BUPs) and best invariant predictors (BIPs) for $Y_{j:m}$ for two cases: μ known and unknown. Another logical approach is to consider $E(Y_{j:m}|\mathbf{R}_{s,n})$ as a predictor of $Y_{j:m}$, i.e.

$$\widehat{Y}_{j:m} = E(Y_{j:m}|\mathbf{R}_{s,n}). \tag{2}$$

It may be noted that if $Z_1, Z_2, \ldots, Z_m$ are independent identically distributed (i.i.d.) random variables from $\mathrm{Exp}(\mu, \sigma)$, then

$$Y_{j:m} \stackrel{d}{=} \mu + \sum_{i=1}^{j} \frac{Z_i}{m-i+1}, \tag{3}$$

where $\stackrel{d}{=}$ stands for identical in distribution. This result was originally due to Sukhatme (1937). Since $Y_{j:m}$ and $\mathbf{R}_{s,n}$ are independent, from (2) and (3) we have

$$E(Y_{j:m}|\mathbf{R}_{s,n}) = E(Y_{j:m}) = \mu + \sigma g(j; m), \tag{4}$$

where

$$g(j; m) = \sum_{i=1}^{j} \frac{1}{m-i+1} = \sum_{i=m-j+1}^{m} \frac{1}{i}. \tag{5}$$

To construct predictors for $Y_{j:m}$ in this case, it suffices to estimate μ and σ on the basis of an information sample, i.e. $\mathbf{R}_{s,n} = (R_s, R_{s+1}, \ldots, R_n)$. With this in mind, in the following sections we consider the maximum likelihood estimator (MLE) for two cases: μ known and unknown. Note that if we consider the best linear unbiased estimator (BLUE) and best linear invariant estimator (BLIE), then the BUP and BIP will be recovered, respectively. In Section 2, it is supposed that two parameters are unknown and we find several point predictors and compare them. In Section 3, under the assumption that μ is known, we also obtain maximum likelihood predictors (MLPs). It is well-known that the total time on test (TTT) plays an important role in the reliability analysis and life testing, so we also derive point predictors for the TTT in each case: μ known and unknown. A simulation study and a real example are given in Section 4 to illustrate the proposed procedures.

2. Point prediction: μ and σ are unknown

In this section, under the assumption that the two parameters μ and σ are unknown we construct several point predictors for the order statistics of the future sample based on observed record values. First we present the BUP and BIP for the jth order statistic of a future sample of size m, i.e. $Y_{j:m}$. According to the approach explained in the introduction, we also plug the MLEs of μ and σ into (4) in order to find another kind of predictor for $Y_{j:m}$. We also consider some special cases such as the minimum and the maximum of a future sample and also the TTT. In each case, we derive the corresponding predictors.

2.1. The best unbiased predictor

First, let us give a brief description of the general results for location–scale families. Suppose that Y is a predictable random variable in future and $\mathbf{X} = (X_1, \ldots, X_k)$ is the vector of the information sample. Suppose further that the joint pdf of Y and $\mathbf{X}$ belongs to a location–scale family, i.e.,

$$f(y, x_1, \ldots, x_k, \mu, \sigma) = \sigma^{-(k+1)} f((y-\mu)/\sigma, (x_1-\mu)/\sigma, \ldots, (x_k-\mu)/\sigma).$$