

# Statistical Inference on the Basis of Sequential Order Statistics under a Linear Trend for Conditional Proportional Hazard Rates

Majid Hashempour<sup>1,\*</sup>, Zohreh Pakdaman<sup>1</sup>, Mahdi Doostparast<sup>2</sup>

<sup>1</sup> Department of Statistics, School of Sciences, University of Hormozgan, Bandar Abbas, Iran

<sup>2</sup> Department of Statistics, School of Mathematical Sciences, Ferdowsi University of Mashhad, Mashhad, Iran

**Abstract** This paper deals with systems consisting of independent and heterogeneous exponential components. Since failures of components may change lifetimes of surviving components because of load sharing, a linear trend for conditionally proportional hazard rates is considered. Estimates of parameters, both point and interval estimates, are derived on the basis of observed component failures for  $s (\geq 2)$  systems. Fisher information matrix of the available data is also obtained which can be used for studying asymptotic behaviour of estimates. The generalized likelihood ratio test is implemented for testing homogeneity of  $s$  systems. Illustrative examples are also given.

**Keywords** Dynamic systems; Linear trend model; Exponential distribution; Hazard function

**AMS 2010 subject classifications** 62G30, 60K10, 62P30, 62N05

**DOI:** 10.19139/soic-2310-5070-802

## 1. Introduction

Let  $X_1, \dots, X_n$  be independent and identically distributed (i.i.d.) random variables with a common distribution function (DF), say  $F$ , and abbreviated by  $X_1, \dots, X_n \stackrel{i.i.d.}{\sim} F$ . Denote in magnitude order of  $X_1, \dots, X_n$  by  $X_{1:n} \leq \dots \leq X_{n:n}$ , which are known as order statistics (OSs). The theory of OSs has been widely studied in literature specially in system reliability analyses. For example, lifetimes of known  $r$ -out-of- $n$  systems coincide to  $X_{r:n}$  where  $X_1, \dots, X_n$  stand for component lifetimes; For more information, see Barlow and Proschan [3], David and Nagaraja [11] and references therein. In order to introduce more flexible models for analysing practical systems, various generalizations of OSs such as *fractional order statistics* and *generalized order statistics* have been proposed. The former is useful for providing more flexible tools and the later is a setting to unify similar results (David and Nagaraja [11], p. 21). In this paper, we deal with another unified concept, called *sequential order statistics* (SOS). There is also another motivation in reliability analyses for implementing SOS. Specifically, when component lifetimes are i.i.d., the OSs are suitable for describing the  $r$ -out-of- $n$  system lifetime. Thus failing a component does not effect the DFs of lifetimes of surviving components. As motivated by Cramer and Kamps [7], in practice the failure of a component may result in a higher load on the surviving components and hence causes lifetime distributions change. This property may be due to load sharing and/or common working environments and hence dependent component lifetimes. More precisely, suppose that  $F_j$  and  $f_j$ , for  $j = 1, \dots, n$ , denote the common DF and probability density function (PDF) of the component lifetimes when  $n - j + 1$  components are jointly working. Then, the components begin to work independently at time  $t = 0$  with the common DF  $F_1$ . When at

\*Correspondence to: Majid Hashempour (Email: ma.hashempour@hormozgan.ac.ir). Department of Statistics, School of Sciences, University of Hormozgan, Bandar Abbas, Iran.

time  $x_1$ , the first component failure occurs, the remaining  $n - 1$  components are working with the (left truncated) common DF  $F_2$  at  $x_1$ . This process continues up to  $n - r + 1$  components with the common DF  $F_r$  work until the  $r$ -th failure occurs at time  $x_r$  and hence the whole system fails. This system is called *sequential r-out-of-n system* (or dynamic system) and the system lifetime is then  $r$ -th observed component failure time, denoted by  $X_{(r)}$ . In the literature,  $(X_{(1)}, \dots, X_{(r)})$  is called SOSs. Let  $\mathbf{x} = (x_1, \dots, x_r)$  be the observed values from SOSs  $(X_{(1)}, \dots, X_{(r)})$  with DFs  $(F_1, \dots, F_r)$  and PDFs  $(f_1, \dots, f_r)$  of component lifetimes. The joint PDF of  $X_{(1)}, \dots, X_{(r)}$  ( $r \leq n$ ) is

$$L(F_1, \dots, F_r; \mathbf{x}) = \left( \frac{n!}{(n-r)!} \right) \left\{ \prod_{j=1}^{r-1} f_j(x_j) \left( \frac{\bar{F}_j(x_j)}{\bar{F}_{j+1}(x_j)} \right)^{n-j} \right\} f_r(x_r) \bar{F}_r(x_r)^{n-r}, \tag{1}$$

where  $x_1 < x_2 < \dots < x_n$  and  $\bar{F}_j(x_j) = 1 - F_j(x_j)$ ,  $j = 1, \dots, r$ . Statistical properties of SOSs have been studied by Kamps [15, 16], Cramer and Kamps [7, 8], Balakrishnan *et al.* [1], Beutner and Kamps [6], Esmailian and Doostparast [12], Bedbur [5], Hashempour and Doostparast [13] and references therein. In Equation (1),  $F_j$  ( $j = 1, \dots, r$ ) depends usually on some unknown parameters. Let the number of independent parameters of  $F_j$  be  $k_j$ . Thus, Equation (1) depends on  $\sum_{j=1}^r k_j$  parameters. We consider a simplified statistical model to reduce

the dimension of the parameter space in Equation (1) which is called proportional hazard rate model.

A large family of models introduced by Cox (1972) focuses directly on the hazard rate function. The simple member of the family is the proportional hazard rate model. Different kinds of proportional hazard models may be obtained by making different assumptions about the baseline survival function, or equivalently, the baseline hazard rate function. Let  $F_0(\cdot)$  be a absolutely continuous DF with a corresponding PDF  $f_0(\cdot)$ . The hazard rate function is defined by  $h_0(t) = f_0(t)/\bar{F}_0(t)$  for  $t > 0$ , where  $\bar{F}_0(t) = 1 - F_0(t)$  is the survival function of the DF  $F_0(\cdot)$ . If  $X$  is a member of proportional hazard family with the baseline DF  $F_0(\cdot)$ , then the survival function of  $X$  becomes  $\bar{F}(t; \theta) = \bar{F}_0^\theta(t)$ ,  $t \in S$ , where  $\theta$  is the proportional parameter and  $F_0(\cdot)$  is the baseline DF and  $S$  is the support of the baseline DF. In this case, the hazard rate function of  $X$  is given by  $h(t; \theta) = f(t; \theta)/\bar{F}(t; \theta) = \theta h_0(t)$  for  $t > 0$ .

In this paper, we consider the problem of estimating the parameters on the basis of  $s$  ( $\geq 2$ ) independent SOSs samples under a proposed linear trend conditional proportional hazard rates (LTCPHR) model, defined by  $\bar{F}_j(t) = \bar{F}_0^{aj}(t)$  for  $j = 1, \dots, r$ , where  $aj = a \times j$ ,  $a > 0$  and  $F_0(t)$  is the underlying DF. Remember that the hazard rate function of the DF  $F$  defined by  $h(t) = f(t)/\bar{F}(t)$  for  $t > 0$ , where  $f(t) = \partial F(t)/\partial t$  is the probability density function (PDF) of the DF  $F(t)$ . Therefore,  $h_j(t) = aj h_0(t)$ , for  $j = 1, \dots, r$ , is the proportional hazard rate function of the DF  $F_j$ , where  $h_0(t)$  is hazard rate function of the baseline DF  $F_0$  for all  $t$ .

The LTCPHR model is a new defined statistical concept for modelling engineering systems in which components share the system load. In fact, impact of failing a component on the surviving components are modeled via hazard rate components. Notice that  $F_i$  is the common component distribution function when  $n - i + 1$  components are jointly working. The connection between  $F_i$  is done by assuming a proportional hazard rate among them. Hence, it is called conditionally proportional hazard rate models.

In this paper, we consider the problem of the estimation parameters of the LTCPHR model with independent multiple SOS samples coming from heterogeneous exponential populations. Thus, this paper is organized as follows: In Section 2, the maximum likelihood estimates (MLEs) of parameters are derived and the generalized likelihood ratio test (GLRT) is used for testing homogeneity of the parent exponential populations. In Section 3, we analyse a simulated data set. Finally, some concluding remarks are given in Section 4.

## 2. Statistical inference for the LTCPHR model parameters

In this section, we obtain MLEs of LTCPHR model parameters. Also, GLRT is derived for testing homogeneity of populations. To do these, two scenarios, namely, (i) the parameter  $a$  is known; and (ii) the parameter  $a$  is unknown, are considered.

**2.1. Maximum likelihood estimation**

Suppose that we observed  $s (\geq 2)$  independent heterogeneous SOS samples. The available data may be represented as

$$\mathbf{x} = [[x_{ij}]]_{i=1, \dots, s, j=1, \dots, r}, \tag{2}$$

where the  $i$ -th row of the matrix  $\mathbf{x}$  in (2) denotes the SOS sample coming from the  $i$ -th system. The likelihood function (LF) of the available data given by (2) is then derived from (4) as

$$L(\mathcal{F}; \mathbf{x}) = \left( \frac{n!}{(n-r)!} \right)^s \prod_{i=1}^s \left( \prod_{j=1}^{r-1} \left[ f_j^{[i]}(x_{ij}) \left( \frac{\bar{F}_j^{[i]}(x_{ij})}{\bar{F}_{j+1}^{[i]}(x_{ij})} \right)^{n-j} \right] f_r^{[i]}(x_{ir}) \bar{F}_r^{[i]}(x_{ir})^{n-r} \right), \tag{3}$$

where  $\mathcal{F} = \{F_j^{[i]}, i = 1, \dots, s, j = 1, \dots, r\}$  and for  $i = 1, \dots, s, j = 1, \dots, r, \bar{F}_j^{[i]}(x) = 1 - F_j^{[i]}(x), F_j^{[i]}$  and  $f_j^{[i]}$  denote the survival function, DF and PDF of  $j$ th component lifetime of the  $i$ -th dynamic system, respectively. For more details, see Cramer and Kamps [8, 9] and Hashempour and Doostparast [13]. Suppose that the baseline DF of the  $i$ -th dynamic system ( $i = 1, \dots, s$ ) follows the exponential distribution with the mean  $\sigma_i$ , i.e.

$$F_0(t) = 1 - \exp\left\{-\frac{t}{\sigma_i}\right\}, \quad t > 0, \quad \sigma_i > 0. \tag{4}$$

It should be mentioned that the baseline DFs for the considered LTCPHR model for component lifetimes are heterogeneous with different scale parameters. Therefore under the earlier mentioned LTCPHR model in Section 1, we have  $\bar{F}_j^{[i]}(x_{ij}) = \exp\{-ajx_{ij}/\sigma_i\}$  and  $f_j^{[i]}(x_{ij}) = aj\sigma_i^{-1} \exp\{-ajx_{ij}/\sigma_i\}$  for  $i = 1, \dots, s$  and  $j = 1, \dots, r$ . Then, Equation(3) yields the LF of the available data as

$$L(\sigma_1, \dots, \sigma_s, a; \mathbf{x}) = \left( \frac{r!n!}{(n-r)!} \right)^s (a)^{rs} \left( \prod_{i=1}^s \sigma_i \right)^{-r} \exp \left\{ - \sum_{i=1}^s \sum_{j=1}^r \left( \frac{x_{ij}m_j}{\sigma_i} \right) \right\}, \tag{5}$$

where  $a > 0$ , and  $m_j = (n-j+1)ja - (n-j)(j+1)a$  with convention  $(n-r)(r+1)a \equiv 0$ . For sake of brevity, we assumed that the proportional parameter  $a$  are the same among the  $s$  sequential  $r$ -out-of- $n$  systems. Following Cramer and Kamps [8] and Hashempour and Doostparast [13], two cases are considered in sequel: (i)  $a$  is known, and (ii)  $a$  is unknown.

**Case I: The parameter  $a$  is known**

Suppose that the parameter  $a$  in Equation (5) is known. If  $\sigma_1 = \dots = \sigma_s$ , the ML estimate of the common mean of the  $s$  baseline exponential populations, say  $\sigma_0$ , is derived by maximizing (5) with respect to  $\sigma_0$  as

$$\begin{aligned} \hat{\sigma}_0 &= \frac{a}{rs} \sum_{i=1}^s \sum_{j=1}^r (n-j+1)jD_{ij} \\ &= \frac{1}{rs} \sum_{i=1}^s \sum_{j=1}^r x_{ij}m_j, \end{aligned} \tag{6}$$

where  $D_{ij} = x_{ij} - x_{i,j-1}$ , for  $j = 1, \dots, r$ , with convention  $x_{i0} := 0$  for  $i = 1, \dots, s$ . Notice that  $\sum_{j=1}^r x_{ij}m_j = \sum_{j=1}^r j(n-j+1)aD_{ij}$ , for  $i = 1, \dots, s$ .

If baseline exponential populations are heterogeneous, the (unique) ML estimate of  $\sigma_i$  ( $i = 1, \dots, s$ ) is derived from Equation (5) as

$$\hat{\sigma}_i = \frac{a}{r} \sum_{j=1}^r (n-j+1)jD_{ij} = \frac{1}{r} \sum_{j=1}^r x_{ij}m_j. \tag{7}$$

Under the LTCPHR with the one-parameter exponential baseline DF, we have

$$Q_i := \sum_{j=1}^r (n - j + 1)jaD_{ij} \sim \Gamma(r, \sigma_i), \quad i = 1, \dots, s, \tag{8}$$

where  $\Gamma(m, n)$  calls for the gamma distribution with the shape and the scale parameters  $m$  and  $n$ , respectively. From Equation (8) and for  $i = 1, \dots, s$ ,  $\hat{\sigma}_i \sim \Gamma(r, \sigma_i/r)$ , and then  $E(\hat{\sigma}_i) = \sigma_i$  and  $Var(\hat{\sigma}_i) = \sigma_i^2/r$ . Notice that the ML estimate  $\hat{\sigma}_0$  in Equation (6) is the arithmetic mean of the ML estimates  $\hat{\sigma}_i$  for the mean populations given by Equation (7), i.e.  $\hat{\sigma}_0 = \sum_{i=1}^s \hat{\sigma}_i/s$ . So,  $E(\hat{\sigma}_0) = \bar{\sigma}$  and  $Var(\hat{\sigma}_0) = \bar{\sigma}^2/sr$  where  $\bar{\sigma} = \sum_{i=1}^s \sigma_i/s$  and  $\bar{\sigma}^2 = \sum_{i=1}^s \sigma_i^2/s$ .

**Case II: The parameter  $a$  is unknown**

Suppose that the parameter  $a$  in Equation (5) is unknown. In this case, calculations are complicated. The logarithm of LF in Equation (5) can be written as

$$l(\sigma_1, \dots, \sigma_s, a; \mathbf{x}) = rs \ln a - r \sum_{i=1}^s \ln \sigma_i - a \sum_{i=1}^s \sum_{j=1}^r \frac{j(n - j + 1)D_{ij}}{\sigma_i}. \tag{9}$$

The ML estimates of the parameters which are shown by  $\hat{\sigma}_i$  and  $\hat{a}$  for  $i = 1, \dots, s$ , (if exist) are obtained (numerically) by solving the following likelihood equations:

$$\frac{\partial l}{\partial a} = \frac{rs}{a} - \sum_{i=1}^s \sum_{j=1}^r \frac{j(n - j + 1)D_{ij}}{\sigma_i} = 0, \tag{10}$$

and

$$\frac{\partial l}{\partial \sigma_k} = \frac{a}{\sigma_k^2} \sum_{j=1}^r j(n - j + 1)D_{kj} - \frac{r}{\sigma_k} = 0, \quad k = 1, \dots, s. \tag{11}$$

From Equations (10) and (11), we have

$$\hat{a}(\sigma_1, \dots, \sigma_{k-1}, \sigma_{k+1}, \dots, \sigma_s) = \frac{r(s - 1)}{\sum_{i=1, i \neq k}^s \sum_{j=1}^r \frac{j(n-j+1)D_{ij}}{\sigma_i}}, \tag{12}$$

and

$$\hat{\sigma}_k(\sigma_1, \dots, \sigma_{k-1}, \sigma_{k+1}, \dots, \sigma_s) = \frac{(s - 1) \sum_{j=1}^r j(n - j + 1)D_{kj}}{\sum_{i=1, i \neq k}^s \sum_{j=1}^r \frac{j(n-j+1)D_{ij}}{\sigma_i}}, \quad k = 1, \dots, s. \tag{13}$$

Equations (12) and (13) cannot be solved analytically. The matrix of second derivatives of the likelihood with respect to the parameters is called Hessian matrix (**HM**), that is  $\mathbf{HM} = [[(\partial^2 \log(L))/\partial \theta_i \partial \theta_j]_{1 \leq i, j \leq s+1}]$ , where  $\theta_i = \sigma_i, (1 \leq i, j \leq s)$  and  $\theta_i = a, (i, j = s + 1)$ . For more information, see Khuri [17]. After some algebraic calculations, we have

$$\mathbf{HM} = \begin{pmatrix} \mathbf{B}_{11} & \mathbf{B}_{12} \\ \mathbf{B}_{21} & \mathbf{B}_{22} \end{pmatrix}, \tag{14}$$

where

$$\begin{aligned}
 \mathbf{B}_{11} &= \left\{ {}_d \frac{r}{\sigma_i^2} - \frac{2a \sum_{k=1}^r k(n-k+1)D_{ij}}{\sigma_i^3} \right\}_{i=1, \dots, s}, \\
 \mathbf{B}_{22} &= \left\{ \frac{-sr}{a^2} \right\}_{j=1, \dots, r}, \\
 \mathbf{B}_{12} &= \mathbf{B}_{21}^T = \left[ \left[ \frac{\sum_{j=1}^r (n-j+1)D_{ij}}{\sigma_i^2} \right] \right]_{i=1, \dots, s, j=1, \dots, r},
 \end{aligned}$$

and  $\{ {}_d a_i \}$  denotes a diagonal matrix with element  $a_i$  on the main diagonal. The Hessian matrix (14) is not necessary negative definite on the parameter space. Therefore, one needs to use numerically methods for maximizing the LF (9) with respect to  $a$  and  $\sigma_1, \dots, \sigma_s$  by using Equations (12) and (13).

**2.2. Generalized likelihood ratio test**

In this section, we consider the problem of homogeneity testing on the basis of independent SOS samples from different exponential populations, i.e.,

$$H_0 : \sigma_1 = \dots = \sigma_s \text{ v.s. } H_1 : \sigma_i \neq \sigma_j \exists i \neq j. \tag{15}$$

**Case I: The parameter  $a$  is known**

The GLRT statistic for testing the problem of hypotheses (15) is

$$\begin{aligned}
 \Lambda_1 &= \frac{\sup_{\Omega_0} L(\sigma_1, \dots, \sigma_s; \mathbf{x})}{\sup_{\Omega} L(\sigma_1, \dots, \sigma_s; \mathbf{x})} \\
 &= \prod_{i=1}^s \left( \frac{\hat{\sigma}_i}{\hat{\sigma}_0} \right)^r \exp \left\{ \sum_{i=1}^s \sum_{j=1}^r \left( \frac{1}{\hat{\sigma}_i} - \frac{1}{\hat{\sigma}_0} \right) m_j x_{ij} \right\},
 \end{aligned} \tag{16}$$

where  $\Omega = \{(\sigma_1, \dots, \sigma_s) : \sigma_i > 0, i = 1, \dots, s\} =: \mathbb{R}^{+s}$  is the whole parameter space and  $\Omega_0 = \{(\sigma_1, \dots, \sigma_s) : \sigma_1 = \dots = \sigma_s, \sigma_i > 0, i = 1, \dots, s\}$  denotes the parameter space under the null hypothesis  $H_0$ . After some algebraic manipulations, the logarithm of the GLRT statistic  $\Lambda_1$  given by Equation (16) reduces to

$$\log \Lambda_1 = r \sum_{i=1}^s \log \left( \frac{sQ_i}{\sum_{j=1}^s Q_j} \right), \tag{17}$$

where  $Q_i$  is defined by Equation (8) and “log” stands for the natural logarithm. Then, the null hypothesis  $H_0$  is rejected if

$$A(\mathbf{Q}, \mathbf{a}) > c, \tag{18}$$

where  $\mathbf{Q} = (Q_1, \dots, Q_s)$  and

$$A(\mathbf{Q}, \mathbf{a}) = - \sum_{i=1}^s \log \left( \frac{Q_i}{\sum_{j=1}^s Q_j} \right).$$

The constant  $c$  in Equation (18) is obtained subject to the level of the test, say  $\gamma$ . To derive the constant  $c$  in Equation (18), we need the distribution of the vector  $\mathbf{Q}$  under the null hypothesis  $H_0$  in (15). To do this, we need the following lemma.

*Lemma 1*

(Balakrishnan and Nevzorov [2])

Let  $Z_1, \dots, Z_k$  be independent random variables and  $Z_i \sim \Gamma(a_i, 1)$ , for  $i = 1, \dots, k$ . Then  $(U_1, \dots, U_k) \sim$

$D(a_1, \dots, a_k)$  where  $U_i = Z_i / \sum_{j=1}^k Z_j$ , for  $i = 1, \dots, k$ . Here,  $D(a_1, \dots, a_k)$  stands for the Dirichlet distribution with PDF

$$f(u_1, \dots, u_k) = \frac{\Gamma(\sum_{i=1}^k a_i)}{\prod_{i=1}^k \Gamma(a_i)} \left( \prod_{i=1}^k u_i^{a_i-1} \right), \quad 0 \leq u_i \leq 1, \quad \sum_{i=1}^k u_i = 1.$$

Let  $V_i = Q_i / \sum_{j=1}^s Q_j$ , for  $i = 1, \dots, s-1$  and  $V_s = -\sum_{i=1}^s \log(Q_i / \sum_{j=1}^s Q_j)$ . The Jacobian transformation is then

$$J = \exp\{-v_s\} / \prod_{i=1}^{s-1} v_i. \tag{19}$$

The joint PDF of  $(V_1, \dots, V_s)$  under the homogeneity hypothesis  $H_0$  in (15) is derived from Equations (8) and (19) and Lemma 1 as

$$\begin{aligned} f_{V_1, \dots, V_s}(v_1, \dots, v_s) &= \frac{\Gamma(sr)}{\Gamma(r)^s} \left( \prod_{i=1}^{s-1} v_i^{r-1} \right) \frac{\exp\{-(r-1)v_s\} \exp\{-v_s\}}{\prod_{i=1}^{s-1} v_i^{r-1} \prod_{i=1}^{s-1} v_i} \\ &= \frac{\Gamma(sr) \exp\{-rv_s\}}{\Gamma(r)^s \prod_{i=1}^{s-1} v_i}, \end{aligned} \tag{20}$$

for  $v_i \geq 0$ ,  $1 \leq i \leq s$  and  $\sum_{i=1}^{s-1} v_i + \exp\{-v_s\} / \prod_{i=1}^{s-1} v_i = 1$ . Therefore, the marginal PDF of  $V_s$  is readily obtained from Equation (20) as

$$f_{V_s}(v_s) = \int \int \dots \int_B \frac{\Gamma(sr) \exp\{-rv_s\}}{\Gamma(r)^s \prod_{i=1}^{s-1} v_i} dv_1 \dots dv_{s-1}, \tag{21}$$

where  $B = \{(v_1, \dots, v_{s-1}) | v_i \geq 0, \sum_{i=1}^{s-1} v_i + \exp\{-v_s\} / \prod_{i=1}^{s-1} v_i = 1\}$ . In practice, one may use numerical methods such as Monte Carlo simulation to derive the threshold  $c$  in the rejection region (18). For more details, see Hashempour [14].

*Remark 1*

It is easy to verify that the distribution family (5) is invariant with respect to the group of the scale transformations

$$\mathcal{G} = \{g_d : g_d(\mathbf{x}) = d\mathbf{x} = \{dx_{ij}^*\}_{1 \leq i \leq s, 1 \leq j \leq r}, d > 0\}. \tag{22}$$

Also, the problem of hypotheses testing (15) remains invariant under  $\mathcal{G}$  in (22) since  $\bar{G}(\Omega) = \Omega$  and  $\bar{G}(\Omega_0) = \Omega_0$  where  $\Omega = \{(\sigma_1, \dots, \sigma_s) : \sigma_i > 0, i = 1, \dots, s\} = \mathbb{R}^{+s}$ ,  $\Omega_0 = \{(\sigma_1, \dots, \sigma_s) : \sigma_1 = \dots = \sigma_s, \sigma_i > 0, i = 1, \dots, s\}$  and  $\bar{G} = \bar{g}_d(\sigma_1, \dots, \sigma_s) = (d\sigma_1, \dots, d\sigma_s)$  is the induced group of transformations by the group of transformations  $\mathcal{G}$  in Equation (22) on the parameter space  $\Omega$ . Finding the uniformly most powerful invariant test (if exists) for the problem (15) remains as an open problem.

From Equation (8), one can see that  $2r(\hat{\sigma}_i / \sigma_i) \sim \chi_{2r}^2$ , where  $\chi_\nu^2$  stands for the chi-square distribution with  $\nu$  degrees of freedom. So, an equi-tailed confidence interval at level  $100\gamma\%$  for  $\sigma_i$  ( $i = 1, \dots, s$ ) is

$$\left( \frac{2r\hat{\sigma}_i}{\chi_{2r, (1+\gamma)/2}^2}, \frac{2r\hat{\sigma}_i}{\chi_{2r, (1-\gamma)/2}^2} \right), \tag{23}$$

where  $\chi_{\nu, p}^2$  calls for the  $p$ -th quantile of the  $\chi_\nu^2$ -distribution.

Note that the observed Fisher Information (FI), denoted by  $I_{jk}(\hat{\sigma}_1, \dots, \hat{\sigma}_s)$ , on the basis of the available SOSs data (2) is equal to minus of the Hessian matrix (HM) at the point of MLEs, i.e.

$$I_{jk}(\hat{\sigma}_1, \dots, \hat{\sigma}_s) = [(-\partial^2 \log(L)/\partial\sigma_i\partial\sigma_j)_{1 \leq i, j \leq s}]|_{\sigma_1=\hat{\sigma}_1, \dots, \sigma_s=\hat{\sigma}_s}.$$

It is well known that the unique MLEs follow asymptotically the multivariate normal distribution with mean vector  $(\sigma_1, \dots, \sigma_s)$  and the variance-covariance matrix  $[I_{jk}(\hat{\sigma}_1, \dots, \hat{\sigma}_s)]^{-1}$  (see, e.g., Lehmann and Romano, [18]). Therefore, an approximate equi-tailed  $100\gamma\%$  confidence interval for  $\sigma_i$  is

$$\left( \hat{\sigma}_i - z_{(1-\gamma)/2} \sqrt{\frac{\hat{\sigma}_i^2}{r}}, \hat{\sigma}_i + z_{(1-\gamma)/2} \sqrt{\frac{\hat{\sigma}_i^2}{r}} \right), \tag{24}$$

where  $z_\gamma$  stands for the  $\gamma$ -quantile of the standard normal distribution.

**Case II: The parameter  $a$  is unknown**

It is easy to verify that the unique ML estimates of the parameters under the null hypothesis  $H_0$  are

$$\hat{\sigma}_0 = \frac{\sum_{i=1}^s \sum_{j=1}^r x_{ij} \hat{m}_{0,j}}{rs} = \frac{\hat{a}_0}{rs} \sum_{i=1}^s \sum_{j=1}^r (n-j+1)j D_{ij}, \tag{25}$$

and

$$\hat{a}_0 = \frac{rs}{\sum_{i=1}^s \sum_{j=1}^r (n-j+1)j D_{ij}} \hat{\sigma}_0, \tag{26}$$

where  $\hat{m}_{0,j} = (n-j+1)j\hat{a}_0 - (n-j)(j+1)\hat{a}_0$ , with convention  $\hat{a}_0(r+1) \equiv 0$ . Therefore, the GLRT statistic for the hypotheses testing problem (15) is

$$\Lambda_2 = \prod_{j=1}^r \left( \frac{\hat{a}_0}{\hat{a}} \right)^{rs} \prod_{i=1}^s \left( \frac{\hat{\sigma}_i}{\hat{\sigma}_0} \right)^r \exp \left\{ \sum_{i=1}^s \sum_{j=1}^r \left( \frac{\hat{m}_j}{\hat{\sigma}_i} - \frac{\hat{m}_{0,j}}{\hat{\sigma}_0} \right) x_{ij} \right\}, \tag{27}$$

where  $\hat{m}_j = (n-j+1)\hat{\alpha}_j - (n-j)\hat{\alpha}_{j+1}$ . The logarithm of the GLRT statistic  $\Lambda_2$  in Equation (27) reads

$$\log \Lambda_2 = rs \log \left( \frac{\hat{a}_0}{\hat{a}} \right) + r \sum_{i=1}^s \log \left( \frac{\hat{\sigma}_i}{\hat{\sigma}_0} \right) + \sum_{i=1}^s \sum_{j=1}^r \left( \frac{\hat{m}_j}{\hat{\sigma}_i} - \frac{\hat{m}_{0,j}}{\hat{\sigma}_0} \right) x_{ij}. \tag{28}$$

The null hypothesis  $H_0$  rejects if

$$-2 \log \Lambda_2 > c, \tag{29}$$

where  $-2 \log \Lambda_2$  follows asymptotically chi-square distribution. Exact distribution of the statistic  $\log \Lambda_2$  in Equation (28) under the null hypothesis  $H_0$  is complicated and we could not obtain an explicit expression. This remains as an open problem. In practice, one may use numerical methods such as the Monte Carlo simulation to derive the threshold  $c$  in the rejection region (29). Bedbur [5] obtained the uniformly most powerful unbiased tests under the conditionally proportional hazard rates model based on multiple homogeneous SOS samples from a common exponential distribution.

**3. An illustrative example**

In order to assess the performance of the derived estimates in the preceding sections a simulation study was conducted. In the case of unknown parameter  $a$ ,  $N = 10^4$  multiple SOS samples were generated for some selected values of parameters. For this purpose, we assume that  $a = 0.9, 1, 1.1$ ,  $n = 10, 20$ ,  $s = 4$ ,  $r = 4, 6, 15$

and  $(\sigma_1, \dots, \sigma_s) = (1, \dots, 1)$ . Here, Bias and MSE stand for estimated bias and mean squared error of estimators. The Bias and MSE of the ML estimates of parameters  $(\sigma_1, \dots, \sigma_s)$  and  $a$  are reported in rows of Table 1 from up to down, respectively.

For  $n = 10, s = 1, a = 1.1$  and  $r = 4$ , the equi-tailed confidence intervals at level 95% for unknown parameters  $\sigma_i$  ( $i = 1, \dots, n$ ) in Equation (23) are (0.4570, 3.6768), (0.4658, 3.7471), (0.4477, 3.6018) and (0.4614, 3.7120), respectively. The length of aforementioned confidence intervals for unknown parameters  $\sigma_i$  for  $i = 1, \dots, s$  are 3.2208, 3.2813, 3.1540 and 3.2506, respectively. Also, from Equation (24), the approximate 95% confidence intervals for the unknown parameters  $\sigma_i$  for  $i = 1, \dots, s$  are given by (0.0200, 1.9835), (0.0200, 1.9835), (0.0200, 1.9835) and (0.0200, 1.9835), respectively.

The length of approximate confidence intervals for unknown parameters  $\sigma_i$  for  $i = 1, \dots, s$  are 1.9635, 2.0010, 1.9234 and 1.9823, respectively.

Table 1. The *MSE* and *Bias* of ML estimates in case where the parameter  $a$  is unknown.

$n$	$r$	$a = 0.9$		$a = 1$		$a = 1.1$	
		<i>Bias</i>	<i>MSE</i>	<i>Bias</i>	<i>MSE</i>	<i>Bias</i>	<i>MSE</i>
10	4	-0.0121	0.0662	-0.0129	0.0541	-0.0114	0.0504
		-0.0259	0.0579	-0.0165	0.0533	-0.0151	0.0537
		-0.0277	0.0538	-0.0209	0.0466	-0.0224	0.0486
		-0.0109	0.0484	-0.0257	0.0449	-0.0164	0.0516
		-0.0699	0.0514	-0.0782	0.0668	-0.0900	0.0828
10	6	-0.0095	0.0436	-0.0247	0.0494	-0.0130	0.0448
		-0.0247	0.0518	-0.0262	0.0531	-0.0241	0.0506
		-0.0245	0.0534	-0.0157	0.0451	-0.0136	0.0469
		-0.0202	0.0455	-0.0127	0.0432	-0.0135	0.0506
		-0.0706	0.0508	-0.0797	0.0652	-0.0855	0.0757
20	4	-0.0186	0.0560	-0.0184	0.0427	-0.0133	0.0479
		-0.0125	0.0514	-0.0172	0.0474	-0.0853	0.0438
		-0.0231	0.0481	-0.0198	0.0433	-0.0185	0.0464
		-0.0159	0.0401	-0.0154	0.0493	-0.0164	0.0411
		-0.0712	0.0500	-0.0808	0.0628	-0.0857	0.0763
20	15	-0.0092	0.0510	-0.0130	0.0412	-0.0247	0.0410
		-0.0141	0.0400	-0.0109	0.0461	-0.0065	0.0421
		-0.0239	0.0426	-0.0204	0.0420	-0.0024	0.0441
		-0.0082	0.0396	-0.0230	0.0430	-0.0116	0.0406
		-0.0797	0.0442	-0.0850	0.0537	-0.0961	0.0633

Empirical evidences from Table 1 are summarized as follow;

- The MSEs are decreasing in  $n$  and  $r$ .
- The Biases are negligible and negative.
- The MSE of the obtained numerical estimator of parameter  $a$  is increasing in  $a$ .

#### 4. Conclusions and further remarks

In this paper, based on independent SOSs coming from heterogeneous exponential populations under a linear trend model, the MLEs parameters were obtained on the basis of multiple SOS samples. The GLR tests were derived for testing homogeneity of the exponential populations. Some open problems were also mentioned. The results of this paper may be extended in some directions. For example, derivation of the uniformly most powerful scale-invariant test (if exist) is worth for further consideration.



## Acknowledgement

The authors are grateful to three anonymous referees for making many helpful comments and suggestions on an earlier version of this paper. This research was supported by a grant from University of Hormozgan.

## REFERENCES

1. N. Balakrishnan, U. Kamps and M. Kateri (2012). A Sequential Order Statistics Approach to Step-Stress Testing. *Annals of the Institute of Statistical Mathematics*, **64**(2), pp. 302–318.
2. N. Balakrishnan and V. B. Nevzorov. (2003). *A Primer on Statistical Distributions*. John Wiley & Sons, Inc., Hoboken, New Jersey.
3. R. E. Barlow and F. Proschan. (1981). *Statistical theory of reliability and life testing: probability models*, Springer, Second Edition.
4. D. R. Cox. (1972). Regression models and life-tables. *Journal of the Royal Statistical Society: Series B (Methodological)*, **34**(2), pp. 187–202.
5. S. Bedbur. (2010). UMPU Tests based on Sequential order statistics. *Journal of Statistical Planning and Inference*, **140**, pp. 2520C-2530.
6. E. Beutner and U. Kamps. (2009). Order restricted statistical inference for scale parameters based on sequential order statistics. *Journal of Statistical Planning and Inference*, **139**, pp. 2963–2969.
7. E. Cramer and U. Kamps. (1996). Sequential order statistics and k-out-of-n systems with sequentially adjusted failure rates. *Annals of the Institute of Statistical Mathematics*, **48**(3), pp. 535–549.
8. E. Cramer and U. Kamps. (2001a). Estimation with Sequential Order Statistics from exponential distributions. *Annals of the Institute of Statistical Mathematics*, **53**(2), pp. 307–324.
9. E. Cramer and U. Kamps. (2001b). Sequential k-out-of-n systems. In N. Balakrishnan and E. Rao, editors, *Handbook of Statistics, Advances in Reliability*, volume **20**, Chapter 12, pp. 301–372.
10. E. Cramer and U. Kamps. (2003). Marginal distributions of sequential and generalized order statistics. *Metrika*, **58**, pp. 293–310.
11. H. A. David and Nagaraja. (2003). *Order Statistics*. John Wiley & Sons, Inc.
12. M. Esmailian and M. Doostparast. (2014). Estimation based on sequential order statistics with random removals. *Probability and Mathematical Statistics*, **34**(1), pp. 81–95.
13. M. Hashempour and M. Doostparast. (2016). Bayesian inference on multiple sequential order statistics from heterogeneous exponential populations with GLR test for homogeneity. *Communications in Statistics-Theory and Methods*, **46**(16), pp. 8086–8100.
14. M. Hashempour. (2017). Classical, Bayesian and evidential inferences based on sequential order statistics, PhD thesis in Mathematical Statistics, Department of Statistics, Ferdowsi University of Mashhad, Mashhad, Iran.
15. U. Kamps. (1995a). *A Concept of Generalized Order Statistics*. Teubner.
16. U. Kamps. (1995b). A concept of generalized order statistics. *Journal of Statistical Planning and Inference*, **48**, pp. 1–23.
17. A. I. Khuri. (2003). *Advanced Calculus With Applications in Statistics*. 2-th Edition, John Wiley & Sons, Inc., Hoboken, New Jersey.
18. E. L. Lehmann, and J. Romano. (2005). *Testing Statistical Hypothesis*. 3-th Edition, Springer, New York.