

## Kullback-Leibler Information in View of an Extended Version of $k$ -Records

Ahmadi Mosayeb<sup>a</sup>, Mohtashami Borzadaran G.R.<sup>1,a</sup>

<sup>a</sup>Department of Statistics, OSDCE, Ferdowsi University of Mashhad

---

### Abstract

This paper introduces an extended version of  $k$ -records. Kullback-Leibler (K-L) information between two generalized distributions arising from  $k$ -records is derived; subsequently, it is shown that K-L information does not depend on the baseline distribution. The behavior of K-L information for order statistics and  $k$ -records, is studied. The exact expressions for K-L information between distributions of order statistics and upper (lower)  $k$ -records are obtained and some special cases are provided.

Keywords: Order statistics,  $k$ -records, Kullback-Leibler information.

---

### 1. Introduction

Statistical analysis is widely used in science to achieve a better insight to the studies of phenomena. Many observational studies lack a well defined sampling frame for the selection of sampling units, particularly in observational studies of humans, wildlife, insects, fisheries and plants. When an investigator registers an observation by nature according to a certain stochastic model, the registered observation will not have the original distribution unless every observation is given an equal chance of being registered. Weighted distribution theory provides a unifying approach for correction of biases that exist in unequally weighted sample data. The concept of weighted distribution, introduced by Fisher (1934) and Rao (1965), has been discussed in many situations. Particular practical views in a wide variety of probability and statistics areas have an important role in statistical inference. Many phenomena in nature can be expressed via the study of the relations between weighted distributions and the parent distributions in various aspects of statistical analysis. We refer the reader to Patil and Rao (1977, 1978) and Rao (1985) for a survey of statistical applications due to weighted distributions (especially data relating to human population ecology). From Gupta and Keating (1986) to Bartoszewicz (2009) and their references, the idea of the weighted distributions aspects and special cases developed into various applications and properties. It is therefore important to study the structure of the distributions due to weighted random variables with respect to the original distribution. Among weights, proportional hazard and proportional reversed hazard models, order statistics, record values and  $k$ -records are those concentrated in ordered data analysis. The obvious question is, “Can we obtain distributions arising extension of the above weights?”. We are interested in concentrating of ordered data such as versions of order statistics and record values.

The origin of order statistics as a case of the ordered data back to Wilks (1948) with many different applications in the area of statistics and its related criteria. After 1948, many books on various aspects of the order statistics have been written such as the book by David and Nagaraja (2003) and Arnold

---

<sup>1</sup> Corresponding author: Department of Statistics, Ordered and Spatial Data Center of Excellence, Ferdowsi University of Mashhad, Mashhad, Iran. E-mail: [grmohtashami@um.ac.ir](mailto:grmohtashami@um.ac.ir)

*et al.* (2008) which are popular books. Many ordered data such as records and  $k$ -records have also progressed and some of their references can be found in Arnold *et al.* (1998); subsequently, numerous papers about this subject have been published. The study of record values is in many ways parallel to order statistics but in a slightly different meaning. Wong and Chen (1990) examined the entropy of ordered sequence of random variables and order statistics. Applying the idea of the K-L information as an extended version of entropy to ordered data cases are focused on here.

Information theory today is considered a very fundamental field that can connect to other areas such as thermodynamics, communication theory, probability theory, statistics and computer science. After Shannon (1948), Kullback and Leibler (1951) defined and studied the properties of measures of information which is now referred to as K-L information measure as an extension of the Shannon entropy. K-L information expresses the distance between any arbitrary probability distribution and a true distribution (like data, observations or a precise distribution). This indicates that it can be used to determine “how far away” a probability distribution is from another distribution. Note that it is not the same as a distance in calculus and we cannot call it a metric on the space of probability distribution. In Bayesian statistics, the K-L information can be used as a measure of information gain in moving from prior to posterior distribution. For more details on information theory see Cover and Thomas (2006). It becomes meaningful when the approximation of the value of K-L information measure is near zero. The K-L information of two distributions near zero means that the two distributions are approximately the same.

This paper presents K-L information for an extended version of  $k$ -records data and its properties for some special cases. The cases of records introduced by MirMostafaei and Ahmadi (2012) and families of distributions arising from order statistics obtained by Jones (2004) are special cases for our model. We have obtained results in view of the K-L information for two models of this generalized version and also achieved various special cases and some of their numerical studies.

The link between order statistics and record data has attracted the attention of many researchers such as Gupta and Ahsanullah (2004), Nevzorov (2000), Gupta (1984), Balakrishnan *et al.* (2009) and Ahmadi and Balakrishnan (2012). As we know, based on Shaked and Shanthikumar (1994, p.496) and Ahmadi and Arghami (2001), one application of records arises in studying the lifetime component with a minimal repairs process; however, the  $k^{th}$  order statistic can be represented as the lifetime of  $(n - k + 1)$ -out-of- $n$  systems. Therefore, finding the information of the lifetime of  $k$ -out-of- $n$  systems can be useful for the prediction of next minimal repair. For more details see, Ahmadi *et al.* (2010). It is necessary to know when we can say the distribution of order statistics and record values are almost identical; subsequently, we ask “Can we approximate records distribution with order statistics distribution and vice versa?” To answer this question, we have applied for two generalized versions of  $k$ -records and concentrate on the analysis of the K-L information between records and order statistics distributions which our achievements here are valuable. The results obtained via the state that “the K-L information between two distributions have amounts of near zero”. A numerical study for special cases and their importance and motivations (the other direction of this paper) are also achieved.

## 2. Preliminaries

As mentioned in Section 1, we are interested in finding the K-L information between extended versions of ordered data. With this in mind, first, we recall some well-known basic definitions and preliminaries.

### 2.1. Order statistics and records

Let  $X_1, X_2, \dots, X_n$  be a random sample of size  $n$  from a continuous population with probability density function (pdf)  $f(x)$  and cumulative distribution function (cdf)  $F(x)$ , and  $X_{1:n} \leq X_{2:n} \leq \dots \leq X_{n:n}$  be the corresponding order statistics. The pdf and cdf of the  $i^{th}$  order statistic i.e.  $X_{i:n}$  ( $1 \leq i \leq n$ ) are given by

$$f_{i:n}(x) = \frac{1}{\beta(i, n-i+1)} F(x)^{i-1} (\bar{F}(x))^{n-i} f(x), \quad (2.1)$$

and

$$F_{i:n}(x) = \sum_{r=i}^n \binom{n}{r} F(x)^r (\bar{F}(x))^{n-r},$$

respectively, where  $\bar{F}(x) = 1 - F(x)$  is the survival function of  $X$  and  $\beta(\cdot, \cdot)$  is the complete beta function. See, David and Nagaraja (2003) or Arnold *et al.* (2008) for more details on the theory and applications of order statistics.

Upper  $k$ -records process is defined in terms of the  $k^{th}$  largest  $X$  yet seen. For a formal definition, see what is given in Arnold *et al.* (1998). In continuous case, let  $T_{1,k} = k$  and for  $n \geq 2$

$$T_{n,k} = \min \left\{ j : j > T_{n-1,k}, X_j > X_{T_{n-1,k}-k+1:T_{n-1,k}} \right\}.$$

The sequence of upper  $k$ -record is then defined by  $R_{n,k}^U = X_{T_{n,k}-k+1:T_{n,k}}$  for  $n \geq 1$ . An analogous definition can be given for  $R_{n,k}^L$ , i.e., lower  $k$ -records as well. The pdf of the  $n^{th}$  upper  $k$ -record and the  $n^{th}$  lower  $k$ -record are given by

$$f_{n,k}(x) = \frac{k^n}{\Gamma(n)} [-\log \bar{F}(x)]^{n-1} (\bar{F}(x))^{k-1} f(x), \quad (2.2)$$

and

$$g_{n,k}(x) = \frac{k^n}{\Gamma(n)} [-\log F(x)]^{n-1} F(x)^{k-1} f(x), \quad (2.3)$$

respectively, where  $\Gamma(\cdot)$  is the complete gamma function. See, Arnold *et al.* (1998) and references therein for more studying on the theory and applications of record data.

### 2.2. Kullback Leibler information

For a convex function  $h : R^+ \rightarrow R$ , the Csiszár  $h$ -divergence between two probability distributions  $f$  and  $g$  is defined as  $C_h(f, g) = \sum_x g(x) h[f(x)/g(x)]$ , by Csiszár (1963). Various information measures are derived via some special values of  $h$ . For example,  $h(x) = x \log x$ ,  $x > 0$  leads to the K-L information as

$$C_h(f, g) = K(f \| g) = \int \log \left( \frac{f(x)}{g(x)} \right) dF(x), \quad (2.4)$$

which is obtained by Kullback and Leibler (1951), where subsume continuous and discrete version. Although it is not a metric, but it sometimes known as K-L distance between  $f$  and  $g$ , with well-known property  $K(f \| g) \geq 0$ , the equality hold if and only if  $f(x) = g(x)$ ,  $\forall x$ . Through this property,

various goodness of fit tests and limit distributions are discussed in many references; therefore, K-L information is a basic tool in decision and information theory with important role in statistical inference. One of the symmetric version of  $K(f \parallel g)$  is given by  $J(f \parallel g) = K(f \parallel g) + K(g \parallel f)$ , which is famous as J-divergence.

### 3. A Generalized Distribution Arising from $k$ -Records Distribution

Ordered data occur in a wide variety with applications in various branches of science and technology. We produce a representation as extended version distribution due to ordered data and introduced the form of its density function. Let  $X$  be a continuous random variable with pdf  $f_\theta(x)$  and cdf  $F_\theta(x)$  with support  $S$  and the vector parameter  $\theta$ ,  $\theta \in \Theta$ (parameter space). We consider an extended version of the  $k$ -record distribution as follows:

$$g_F(x, \eta) = \frac{1}{M_\lambda} [F_\theta(x)]^{a-1} [\bar{F}_\theta(x)]^{b-1} [-\log F_\theta(x)]^{c-1} [-\log \bar{F}_\theta(x)]^{d-1} f_\theta(x), \quad x \in S, \theta \in \Theta, \quad (3.1)$$

where  $M_\lambda$  is normalizing constant and is given by

$$M_\lambda = \int_0^1 u^{a-1} (1-u)^{b-1} (-\log u)^{c-1} [-\log(1-u)]^{d-1} du,$$

such that  $a, b, c$  and  $d$  are positive real constants,  $\eta = (a, b, c, d, \theta)$  and  $\lambda = (a, b, c, d)$ . Model (3.1) as extended version of the  $k$ -record distribution includes many famous families of distributions, such as distribution of order statistics, upper and lower  $k$ -records. Taking various values of  $a, b, c$  and  $d$  in (3.1) implies many common distributions as below:

- Take  $c = d = 1$ , leads to the Jones (2004) model, in this case for  $a = i$  and  $b = n - i + 1$ ,  $g_F(x)$  is the pdf of the  $i^{th}$  order statistic in random sample of size  $n$ .
- Take  $a = b = c = 1$  and  $d = m$ , then  $g_F(x)$  is the same as pdf of the  $m^{th}$  upper record, and for  $a = b = d = 1$  and  $c = n$ ,  $g_F(x)$  is the pdf of the  $n^{th}$  lower record.
- Take  $a = c = 1$ ,  $b = k$  and  $d = m$ , then  $g_F(x)$  is distribution of the  $m^{th}$  upper  $k$ -record, and if  $b = d = 1$ ,  $a = k$  and  $c = n$ , then  $g_F(x)$  is distribution of the  $n^{th}$  lower  $k$ -record.
- Take  $a = b = 1$ , implies MirMostafae and Ahmadi (2012) model.
- Take  $b = c = d = 1$  and  $a = c = d = 1$  are models arising from reversed proportional hazard and proportional hazard, respectively.

It is clear that model with pdf (3.1) subsume many cases as mentioned before. Now, we are interested in properties of these models in view of the K-L information.

### 4. K-L Information between Two Generalized Distribution Arising from $k$ -Record Distribution

Let  $g_F(x)$  and  $g'_F(x)$  be two pdfs from (3.1) with,  $\lambda = (a, b, c, d)$  and  $\lambda' = (a', b', c', d')$ , respectively, then, it can be shown that

$$K[g_F(x) \parallel g'_F(x)] = \log \left( \frac{M_{\lambda'}}{M_\lambda} \right) + (a - a') [\log M_\lambda]_a, \quad (4.1)$$

where  $[\log M_\lambda]_\alpha = \partial \log M_\lambda / \partial \alpha$ . It is observed that K-L information between two pdfs version from (3.1) does not depend on the baseline distribution  $F$ . In subsequent, we obtain exact explicit expressions for (4.1) via some special cases. First, let  $c = d = 1$  and

$$g_i(x) = \frac{1}{\beta(a_i, b_i)} [F(x)]^{a_i-1} [\bar{F}(x)]^{b_i-1} f(x), \quad a_i, b_i \in N = \{1, 2, 3, \dots\},$$

which is a version of Jones (2004) model. Then, from (2.4) we have

$$\begin{aligned} K[g_1(x) \| g_2(x)] &= \int_{-\infty}^{\infty} g_1(x) \log \frac{g_1(x)}{g_2(x)} dx \\ &= \log \left[ \frac{\beta(a_2, b_2)}{\beta(a_1, b_1)} \right] \\ &\quad + \frac{a_1 - a_2}{\beta(a_1, b_1)} \int_0^1 \log(u) u^{a_1-1} (1-u)^{b_1-1} du \\ &\quad + \frac{b_1 - b_2}{\beta(a_1, b_1)} \int_0^1 \log(1-u) u^{a_1-1} (1-u)^{b_1-1} du. \end{aligned} \quad (4.2)$$

However, it is easy to verify that

$$\begin{aligned} [\log M_\lambda]_{a_1} &= \int_0^1 \log(u) u^{a_1-1} (1-u)^{b_1-1} du \\ &= \sum_{r=0}^{b_1-1} \binom{b_1-1}{r} \frac{(-1)^{r+1}}{(a_1+r)^2}, \end{aligned} \quad (4.3)$$

where  $\binom{n}{r} = n! / \{r!(n-r)!\}$ . By substituting (4.3) into (4.2), we obtain

$$K(g_1(x) \| g_2(x)) = \log \left[ \frac{\beta(a_2, b_2)}{\beta(a_1, b_1)} \right] - \frac{a_1 - a_2}{\beta(a_1, b_1)} \sum_{r=0}^{b_1-1} \frac{\binom{b_1-1}{r} (-1)^r}{(a_1+r)^2} - \frac{b_1 - b_2}{\beta(a_1, b_1)} \sum_{r=0}^{a_1-1} \frac{\binom{a_1-1}{r} (-1)^r}{(a_1+b_1+r)^2}. \quad (4.4)$$

**Remark 1.** Note that  $a = a', b = b', c = c'$  and  $d = d'$  deals to various values of K-L information that is mentioned in (4.1), but  $a = a' = 1, b = b' = 1, c = c' = 1$  and  $d = d' = 1$  also implies the corresponding values of K-L information as mentioned before, which are models arising from weighted upper  $k$ -record, weighted lower  $k$ -record, upper  $k$ -record and lower  $k$ -record, respectively.

#### 4.1. K-L information between distributions of two order statistics

Let  $X_{i:n}$  be the  $i^{\text{th}}$ ,  $1 \leq i \leq n$ , order statistic in a random sample of size  $n$  from a population with continuous cdf  $F(x)$  and pdf  $f(x)$ . Independently of  $X_{i:n}$ , let  $Y_{j:m}$  be the  $j^{\text{th}}$ ,  $1 \leq j \leq m$  order statistic from a random sample of size  $m$  from the same distribution. Then, by taking  $a_1 = i, b_1 = n - i + 1, a_2 = j$  and  $b_2 = m - j + 1$  in (4.4), which is equivalent to  $\lambda = (i, n - i + 1, 0, 0)$  and  $\lambda' = (j, m - j + 1, 0, 0)$ , we have

$$\begin{aligned} K(f_{i:n} \| f_{j:m}) &= \log \left[ \frac{\beta(j, m - j + 1)}{\beta(i, n - i + 1)} \right] \\ &\quad + \frac{j - i}{\beta(i, n - i + 1)} \sum_{r=0}^{n-i} \frac{\binom{n-i}{r} (-1)^r}{(r+i)^2} - \frac{n - m + j - i}{\beta(i, n - i + 1)} \sum_{r=0}^{i-1} \frac{\binom{i-1}{r} (-1)^r}{(n - i + r + 1)^2}, \end{aligned} \quad (4.5)$$

where  $f_{i:n}$  and  $f_{j:m}$  are pdfs of  $X_{i:n}$  and  $Y_{j:m}$ , respectively. For the case of  $m = n$ , or  $X_{i:n}$  and  $X_{j:n}$  be two order statistics from one random sample  $X_1, \dots, X_n$ , the expression in (4.5) simplifies to the result of Ebrahimi *et al.* (2004), which is

$$K(f_{i:n} \parallel f_{j:n}) = \log \left[ \frac{\Gamma(j)\Gamma(n-j+1)}{\Gamma(i)\Gamma(n-i+1)} \right] - (i-1) \left( \psi(i) - \psi(n-i) - \frac{i-j}{n-i} \right),$$

where  $\psi(\alpha) = d/d\alpha \log \Gamma(\alpha)$  is the digamma function. Consequently, for two adjacent order statistics, we find

$$K(f_{i+1:n} \parallel f_{i:n}) = \frac{1}{i} + \log \frac{n-i}{i} - \psi(n-i) + \psi(i),$$

and

$$K(f_{i:n} \parallel f_{i+1:n}) = \frac{1}{n-i} - \log \frac{n-i}{i} + \psi(n-i) - \psi(i).$$

These two later measures are decreasing for  $i \leq (n+1)/2$  and increasing for  $i \geq (n+1)/2$ , with respect to  $i$ .

#### 4.2. K-L information between distributions of the $m^{th}$ and the $n^{th}$ upper $k$ -records

Let  $\{X_i, i \geq 1\}$  be a sequence of independent and identically distributed (iid) continuous random variables with a cdf  $F(x)$  and pdf  $f(x)$ . Denote the pdf of the  $m^{th}$  and  $n^{th}$  upper  $k$ -records by  $f_{m,k}$  and  $f_{n,k}$ , respectively. Then, from (2.2) and (2.4), we have

$$\begin{aligned} K(f_{n,k} \parallel f_{m,k}) &= \int_{-\infty}^{\infty} f_{n,k}(x) \log \frac{f_{n,k}(x)}{f_{m,k}(x)} dx \\ &= \frac{(n-m)k^n}{\Gamma(n)} \int_{-\infty}^{\infty} (-\log \bar{F}(x))^{n-1} \log(-\log \bar{F}(x)) (\bar{F}(x))^{k-1} f(x) dx + \log \frac{k^{n-m}\Gamma(m)}{\Gamma(n)}. \end{aligned} \quad (4.6)$$

Transformation variable  $y = -\log \bar{F}(x)$ , leads to

$$\begin{aligned} \int_{-\infty}^{\infty} (-\log \bar{F}(x))^{n-1} \log(-\log \bar{F}(x)) (\bar{F}(x))^{k-1} f(x) dx &= \int_0^{\infty} y^{n-1} \log(y) e^{-ky} dy \\ &= \frac{\Gamma(n)}{k^n} [\psi(n) - \log k]. \end{aligned} \quad (4.7)$$

With substituting (4.7) in (4.6), we obtain

$$K(f_{n,k} \parallel f_{m,k}) = \log \frac{k^{n-m}\Gamma(m)}{\Gamma(n)} + (n-m) [\psi(n) - \log k]. \quad (4.8)$$

For the case of  $n = m + j$ , from (4.8), we have

$$K(f_{m+j,k} \parallel f_{m,k}) = j\psi(m+j) + \log \left[ \frac{\Gamma(m)}{\Gamma(m+j)} \right].$$

It can be shown that  $K(f_{m+j,k} \parallel f_{m,k})$  is an increasing function with respect to (w.r.t.)  $j$ .

For two consecutive upper  $k$ -records, we have the following simple expression

$$K(f_{m+1,k} \parallel f_{m,k}) = \psi(m+1) - \log m,$$

which is a decreasing function w.r.t.  $m$ , because

$$\begin{aligned} K(f_{m+1,k} \parallel f_{m,k}) - K(f_{m,k} \parallel f_{m-1,k}) &= \psi(m+1) - \log m - \psi(m) + \log(m-1) \\ &= \log \frac{m-1}{m} + \psi(m+1) - \psi(m) \\ &= \log \frac{m-1}{m} + \frac{1}{m} \leq 0. \end{aligned}$$

For  $k = 1$  from (4.8), we find K-L information between distribution of two usual upper records. Let  $f_m$  and  $f_n$  be the pdf of the usual  $m^{\text{th}}$  and  $n^{\text{th}}$  upper records, respectively, then

$$K(f_n \parallel f_m) = (n-m)\psi(n) + \log \frac{\Gamma(m)}{\Gamma(n)}.$$

**Remark 2.** The K-L information between distributions of two consecutive upper  $k$ -records does not depend on  $k$  and equal to K-L information between distributions of two consecutive usual upper records.

#### 4.3. K-L information between distributions of the $m^{\text{th}}$ upper $k$ -record and the $m^{\text{th}}$ upper $l$ -record

In this subsection, we obtain K-L information between distributions of the  $m^{\text{th}}$  upper  $k$ -record and the  $m^{\text{th}}$  upper  $l$ -record. From (2.2) and (2.4), we have

$$\begin{aligned} K(f_{m,k} \parallel f_{m,l}) &= \int_{-\infty}^{\infty} \frac{k^m (-\log \bar{F}(x))^{m-1}}{\Gamma(m)} \bar{F}(x)^{k-1} f(x) \log \left[ \left( \frac{k}{l} \right)^m \bar{F}(x)^{k-l} \right] dx \\ &= m \log \left( \frac{k}{l} \right) - \frac{(-k)^m (k-l)}{\Gamma(m)} \int_{-\infty}^{\infty} (\log \bar{F}(x))^m \bar{F}(x)^{k-1} f(x) dx \\ &= m \log \left( \frac{k}{l} \right) - \frac{(-k)^m (k-l)}{\Gamma(m)} \int_0^1 (\log(1-u))^m (1-u)^{k-1} du \\ &= m \log \left( \frac{k}{l} \right) + \frac{m(l-k)}{k}. \end{aligned} \tag{4.9}$$

Putting  $l = k + j$  in (4.9) leads to

$$K(f_{m,k+j} \parallel f_{m,k}) = m \log \frac{k+j}{k} - \frac{mj}{k+j},$$

which is an increasing function w.r.t.  $j$ . For two consecutive of the  $m^{\text{th}}$  upper  $k$ -records, we obtain

$$K(f_{m,k+1} \parallel f_{m,k}) = m \log \frac{k+1}{k} - \frac{m}{k+1},$$

where is decreasing function of  $k$  because,

$$\begin{aligned} K(f_{m,k+2} \parallel f_{m,k+1}) - K(f_{m,k+1} \parallel f_{m,k}) &= m \left[ \log \frac{k+2}{k+1} - \frac{1}{k+2} - \log \frac{k+1}{k} - \frac{1}{k+1} \right] \\ &= m \left[ \log \frac{k(k+2)}{(k+1)^2} - \frac{1}{(k+1)(k+2)} \right] < 0, \end{aligned}$$

for  $k \geq 1$ .

**Remark 3.** By proceeding as above, we can obtain similar results for lower  $k$ -records.

## 5. K-L Information between Distributions of Order Statistics and $k$ -Records

Sometimes, we ask “Is it possible to compare the distributions of order statistic and record statistic?” “And in which case two distributions are the same?” For this reason, we first try to find K-L information between distributions of the upper  $k$ -record and the order statistic.

Let  $f_{j:n}$  be the pdf of the  $j^{\text{th}}$  order statistic in sample of size  $n$  from continuous distribution with cdf  $F(x)$  and pdf  $f(x)$ . Independently, let  $f_{m,k}$  be the pdf of the  $m^{\text{th}}$  upper  $k$ -record of a random sequence from the same distribution. Then, from (2.1), (2.2) and (2.4) we have

$$\begin{aligned} K(f_{j:n} \parallel f_{m,k}) &= \int f_{j:n}(x) \log \frac{f_{j:n}(x)}{f_{m,k}(x)} dx \\ &= \log \left[ \frac{\Gamma(m)}{\beta(j, n-j+1)k^m} \right] + \frac{j-1}{\beta(j, n-j+1)} \int_0^1 \log(u) u^{j-1} (1-u)^{n-j} du \\ &\quad + \frac{n-j-k+1}{\beta(j, n-j+1)} \int_0^1 u^{j-1} (1-u)^{n-j} \log(1-u) du \\ &\quad - \frac{m-1}{\beta(j, n-j+1)} \int_0^1 u^{j-1} (1-u)^{n-j} \log[-\log(1-u)] du. \end{aligned}$$

By using (4.3), (4.7) and also  $\int_0^\infty \log(x) e^{-\beta x} dx = (-\gamma - \log \beta)/\beta$ , and doing some algebraic simplification, we obtain

$$\begin{aligned} K(f_{j:n} \parallel f_{m,k}) &= \log \left[ \frac{\Gamma(m)}{\beta(j, n-j+1)k^m} \right] + \frac{1-j}{\beta(j, n-j+1)} \sum_{r=0}^{n-j} \frac{\binom{n-j}{r} (-1)^r}{(r+j)^2} \\ &\quad + \frac{j-n+k-1}{\beta(j, n-j+1)} \sum_{r=0}^{j-1} \frac{\binom{j-1}{r} (-1)^r}{(n-j+r+1)^2} \\ &\quad + (m-1)\gamma + \frac{m-1}{\beta(j, n-j+1)} \sum_{r=0}^{j-1} \frac{\binom{j-1}{r} (-1)^r \log(n-j+r+1)}{n-j+r+1}, \end{aligned} \quad (5.1)$$

where  $\gamma$  is the Euler constant. Now, we arrive in comparing distributions of order statistics and upper records.

### 5.1. Behavior of $K(f_{j:n} \parallel f_{m,1})$

For  $k = 1$  from (5.1), we obtain K-L information between distributions of the  $j^{\text{th}}$  order statistic in



random sample of size  $n$  and the  $m^{th}$  upper record.

$$\begin{aligned} K(f_{j:n} \parallel f_{m,1}) &= \log \left[ \frac{\Gamma(m)}{\beta(j, n-j+1)} \right] + \frac{1-j}{\beta(j, n-j+1)} \sum_{r=0}^{n-j} \binom{n-j}{r} \frac{(-1)^r}{(r+j)^2} \\ &\quad + (m-1)\gamma + \frac{m-1}{\beta(j, n-j+1)} \sum_{r=0}^{j-1} \frac{\binom{j-1}{r} (-1)^r \log(n-j+r+1)}{n-j+r+1} \\ &\quad + \frac{(-n)}{\beta(j, n-j+1)} \sum_{r=0}^{j-1} \binom{j-1}{r} \frac{(-1)^r}{(n-j+r+1)^2}. \end{aligned}$$

It is an increasing function in  $m$ , for  $m \geq 2$ , because

$$\begin{aligned} K(f_{j:n} \parallel f_{m+1,1}) - K(f_{j:n} \parallel f_{m,1}) &= \int f_{j:n}(x) \log \frac{f_{m,1}(x)}{f_{m+1,1}(x)} dx \\ &= \int_0^1 f_{j:n}(u) \log \left[ \frac{m, 1}{-\log(1-u)} \right] du \\ &= \log m - \int_0^1 f_{j:n}(u) \log[-\log(1-u)] du \\ &= \log m - E[\log(-\log(1-U_{j:n}))] \\ &\geq \log 2 + n\gamma > 0. \end{aligned}$$

The inequality is implied by  $\log m \geq \log 2$ , and the fact that  $U_{j:n}$  is smaller than  $U_{n:n}$  with probability one, and so

$$\begin{aligned} E[\log(-\log(1-U_{j:n}))] &\leq E[\log(-\log(1-U_{n:n}))] \\ &= \int_0^1 nu^{n-1} \log[-\log(1-u)] du \\ &\leq n \int_0^1 \log[-\log(1-u)] du = -n\gamma. \end{aligned}$$

In addition,  $K(f_{j:n} \parallel f_{m,1})$  is a decreasing function with respect to  $j$ , for  $m > 1$  and  $j < n$ . We have computed the numerical values of  $K(f_{j:n} \parallel f_{m,1})$  for  $n = 5, 10, 15$  and some selected values of  $j$  and  $m$ , these are presented in Table 1.

From Table 1, we observe that K-L information is increasing in  $m$  and decreasing in  $j$  for  $m > 1$ . For  $m = 1$ , the  $K(f_{j:n} \parallel f_{1,1})$  has a symmetric property w.r.t.  $j$  for fixed  $n$ . It may be noted that Ebrahimi *et al.* (2004) obtained an exact expression for  $K(f_{j:n} \parallel f_{1,1})$ , namely, K-L information between distribution of order statistics and data distribution.

The average of the K-L information between distributions of  $X_{j:n}$  and  $R_{m,1}^U$  is another problem in here. We know that

$$K(f_{j:n} \parallel f_{m,1}) = -H(X_{j:n}) - \int f_{j:n}(x) \log[f_{m,1}(x)] dx,$$

Table 1: Numerical values of  $K(f_{j:n} \| f_{m,1})$ .

$n$	$m$	$j$											
		1	2	3	4	5	6	7	8	9	10	11	12
5	1	0.8090	<b>0.3620</b>	<b>0.2680</b>	<b>0.3620</b>	0.8090	-	-	-	-	-	-	-
	2	2.9960	<b>1.4330</b>	<b>0.6934</b>	<b>0.2555</b>	<b>0.1300</b>	-	-	-	-	-	-	-
	3	5.8760	3.1970	<b>1.8120</b>	<b>0.8418</b>	<b>0.1300</b>	-	-	-	-	-	-	-
	4	9.1610	5.3670	3.3360	<b>1.8330</b>	<b>0.5500</b>	-	-	-	-	-	-	-
	5	12.7300	7.8240	5.1480	3.1130	<b>1.2450</b>	-	-	-	-	-	-	-
10	1	1.4020	0.8820	0.6750	<b>0.5720</b>	<b>0.5280</b>	<b>0.5280</b>	<b>0.5720</b>	0.6750	0.8820	1.4020	-	-
	2	4.2820	2.7080	<b>1.9430</b>	<b>1.4410</b>	<b>1.0720</b>	<b>0.7856</b>	<b>0.5606</b>	<b>0.3917</b>	<b>0.2974</b>	<b>0.3990</b>	-	-
	3	7.8550	5.2270	3.9030	3.0020	2.3090	<b>1.7360</b>	<b>1.2420</b>	<b>0.8012</b>	<b>0.4060</b>	<b>0.9940</b>	-	-
	4	11.8300	8.1520	6.2690	4.9690	3.9510	3.0920	2.3280	<b>1.6160</b>	<b>0.9200</b>	<b>0.2910</b>	-	-
	5	16.1000	11.3600	8.9220	7.2230	5.8810	4.7360	3.7030	2.7190	<b>1.7220</b>	<b>1.6880</b>	-	-
	6	20.5900	14.8000	11.8000	9.7010	8.0340	6.6030	5.3000	4.0440	2.7470	2.4800	-	-
	7	25.2600	18.4200	14.8600	12.3600	10.3700	8.6520	7.0800	5.5530	3.9540	3.1800	-	-
15	1	1.7750	1.2340	1.0020	0.8680	0.7850	0.7320	<b>0.7100</b>	<b>0.6900</b>	<b>0.7100</b>	0.7320	0.7850	0.8680
	2	5.0600	3.4840	2.7160	2.2090	<b>1.8320</b>	<b>1.5340</b>	<b>1.2880</b>	<b>1.0810</b>	<b>0.9041</b>	<b>0.7527</b>	<b>0.6252</b>	<b>0.5228</b>
	3	9.0380	6.4280	5.1220	4.2430	3.5730	3.0290	2.5670	2.1610	<b>1.7980</b>	<b>1.4670</b>	<b>1.1590</b>	<b>0.8702</b>
	4	13.4200	9.7770	7.9340	6.6810	5.7200	4.9300	4.2500	3.6470	3.0980	2.5860	2.0980	<b>1.6230</b>
	5	18.0900	13.4100	11.0300	9.4080	8.1540	7.1180	6.2220	5.4210	4.6850	3.9930	3.3250	2.6630
	6	22.9900	17.2700	14.3600	12.3600	10.8100	9.5290	8.4160	7.4180	6.4960	5.6230	4.7750	3.9270
	7	28.0700	21.3200	17.8600	15.4900	13.6500	12.1200	10.7900	9.5970	8.4880	7.4350	6.4070	5.3730

then

$$\begin{aligned}
\sum_{j=1}^n K(f_{j:n} \| f_{m,1}) &= - \sum_{j=1}^n H(X_{j:n}) - \sum_{j=1}^n \int f_{j:n}(x) \log[f_{m,1}(x)] dx \\
&= -n\bar{H}(X) - \sum_{j=1}^n \int f_{j:n}(x) \log[f_{m,1}(x)] dx \\
&= -n\bar{H}(X) - \sum_{j=1}^n \int_0^1 \frac{u^{j-1}(1-u)^{n-j}}{\beta(j, n-j+1)} \log \left[ \frac{(-\log(1-u))^{m-1}}{\Gamma(m)} f(F^{-1}(u)) \right] du \\
&= -n\bar{H}(X) - \int_0^1 \log \left[ \frac{(-\log(1-u))^{m-1}}{\Gamma(m)} f(F^{-1}(u)) \right] \sum_{j=1}^n j \binom{n}{j} u^{j-1} (1-u)^{n-j} du \\
&= -n\bar{H}(X) + n(m-1)\gamma + n \log \Gamma(m) - n \int_0^1 \log f(F^{-1}(u)) du.
\end{aligned}$$

Hence,

$$\bar{K}(f_{j:n} \| f_{m,1}) = -\bar{H}(X) + (m-1)\gamma + \log \Gamma(m) - \int_0^1 \log f(F^{-1}(u)) du,$$

where  $\bar{H}(X) = 1/n \sum_{i=1}^n H(X_{i:n})$  and  $H(X)$  is entropy of  $X$ .

## 5.2. K-L information between distributions of $X_{j:n}$ and $R_{m,1}^L$

We obtain K-L information between distributions of the  $j^{\text{th}}$  order statistic in random sample of size  $n$

Table 2: Numerical values of  $K(f_{j:n} \| g_{m,1})$ .

$n$	$m$	$j$											
		1	2	3	4	5	6	7	8	9	10	11	12
5	1	0.8090	<b>0.3620</b>	<b>0.2680</b>	<b>0.3620</b>	0.8090	-	-	-	-	-	-	-
	2	<b>0.1193</b>	<b>0.2555</b>	<b>0.6934</b>	<b>1.4330</b>	2.9950	-	-	-	-	-	-	-
	3	<b>0.1223</b>	<b>0.8418</b>	<b>1.8120</b>	3.1970	5.8740	-	-	-	-	-	-	-
	4	<b>0.5307</b>	<b>1.8330</b>	3.3360	5.3670	9.1600	-	-	-	-	-	-	-
	5	<b>1.2270</b>	3.1130	5.1480	7.8240	12.7300	-	-	-	-	-	-	-
10	1	1.4020	0.8820	0.6750	<b>0.5720</b>	<b>0.5280</b>	<b>0.5720</b>	0.6750	0.8820	1.4020	-	-	-
	2	<b>0.4127</b>	<b>0.2974</b>	<b>0.3917</b>	<b>0.5606</b>	<b>0.7856</b>	<b>1.0720</b>	<b>1.4410</b>	<b>1.9430</b>	2.7080	4.2810	-	-
	3	<b>0.1160</b>	<b>0.4060</b>	<b>0.8012</b>	<b>1.2420</b>	<b>1.7360</b>	2.3090	3.0020	3.9030	5.2270	7.8530	-	-
	4	<b>0.2247</b>	<b>0.9200</b>	<b>1.6160</b>	2.3280	3.0920	3.9510	4.9690	6.2690	8.1520	11.8300	-	-
	5	<b>0.6211</b>	<b>1.7220</b>	2.7190	3.7030	4.7360	5.8810	7.2230	8.9220	11.3600	16.1000	-	-
	6	<b>1.2410</b>	2.7470	4.0440	5.3000	6.6030	8.0340	9.7010	11.8000	14.8000	20.5800	-	-
	7	2.0430	3.9540	5.5530	7.0800	8.6520	10.3700	12.3600	14.8600	18.4207	25.2600	-	-
15	1	1.7750	1.2340	1.0020	0.8680	0.7850	0.7320	<b>0.7060</b>	<b>0.6890</b>	<b>0.7060</b>	0.7320	0.7850	0.8680
	2	<b>0.6420</b>	<b>0.3453</b>	<b>0.4537</b>	<b>0.5228</b>	<b>0.6252</b>	<b>0.7527</b>	<b>0.9041</b>	<b>1.0810</b>	<b>1.2880</b>	<b>1.5340</b>	<b>1.8320</b>	2.2090
	3	<b>0.2025</b>	<b>0.3500</b>	<b>0.5983</b>	<b>0.8702</b>	<b>1.1590</b>	<b>1.4670</b>	<b>1.7980</b>	2.1610	2.5670	3.0290	3.5730	4.2430
	4	<b>0.1684</b>	<b>0.6603</b>	<b>1.1480</b>	<b>1.6230</b>	2.0980	2.5860	3.0980	3.6470	4.2500	4.9300	5.7200	6.6810
	5	<b>0.4220</b>	<b>1.2580</b>	<b>1.9860</b>	2.6630	3.3250	3.9930	4.6850	5.4210	6.2220	7.1180	8.1540	9.4080
	6	<b>0.8988</b>	2.0790	3.0470	3.9270	4.7750	5.6230	6.4960	7.4180	8.4160	9.5290	10.8100	12.3600
	7	<b>1.5580</b>	3.0830	4.2900	5.3730	6.4070	7.4350	8.4880	9.5970	10.7900	12.1200	13.6500	15.4900

and the  $m^{th}$  lower record. Let  $g_{m,1}$  be the pdf of the  $m^{th}$  lower record, then we have

$$\begin{aligned}
K(f_{j:n} \| g_{m,1}) &= \int f_{j:n}(x) \log \frac{f_{j:n}(x)}{g_{m,1}(x)} dx \\
&= \log \left[ \frac{\Gamma(m)}{\beta(j, n-j+1)} \right] + \frac{1-j}{\beta(j, n-j+1)} \sum_{r=0}^{n-j} \binom{n-j}{r} \frac{(-1)^r}{(r+j)^2} \\
&\quad + \frac{j-n}{\beta(j, n-j+1)} \sum_{r=0}^{j-1} \binom{j-1}{r} \frac{(-1)^r}{(n-j+r+1)^2} \\
&\quad + (m-1)\gamma + \frac{(m-1)}{\beta(j, n-j+1)} \sum_{r=0}^{n-j} \binom{n-j}{r} (-1)^r \frac{\log(j+r)}{j+r}.
\end{aligned}$$

It is seen that  $K(f_{j:n} \| f_{m,1}) = K(f_{n-j+1:n} \| g_{m,1})$ . It can be shown that  $K(f_{j:n} \| g_{m,1})$  is an increasing function with respect to  $j$  and  $m$ , for  $j > 1$  and  $m > 1$ .

We have computed the numerical values of K-L information between distributions of  $X_{j:n}$  and  $R_{m,1}^L$  for  $n = 5, 10, 15$  and some selected values of  $j$  and  $m$ , these are presented in Table 2.

From Table 2, we observe that K-L information is increasing in  $m$  and  $j$ , for  $m > 1$ . For  $m = 1$ , the  $K(f_{j:n} \| g_{1,1})$  has a symmetric property with respect to  $j$  for fixed  $n$

## 6. Summary

In this paper, the extended version of the  $k$ -records and its relations with Jones (2004) and Mir-Mostafaei and Ahmadi (2012) models are explained. In addition, K-L information between two generalized distributions arising from  $k$ -records and special cases are derived with their achievements. Most of our concentrations focused on the K-L information between distributions of “ $X_{j:n}$  and  $X_{j:m}$ ”, “ $R_{m,k}^U$  and  $R_{n,k}^U$ ”, “ $R_{m,k}^U$  and  $R_{m,l}^U$ ”, “ $X_{j:n}$  and  $R_{m,k}^U$ ” and “ $X_{j:n}$  and  $R_{m,1}^L$ ” and also some numerical studies.

Table 1 and Table 2 show that the distribution of the lower record is closer to the distribution of the lower order statistic and the distribution of the upper record is closer to the distribution of the upper order statistic as expected.

## Acknowledgements

The authors thank the Editor-in-Chief and the referees for their useful comments and constructive criticisms on the original version of this manuscript, which led to this considerably improved version.

## References

- Ahmadi, J. and Arghami, N. R. (2001). Some univariate stochastic orders on record values, *Communication Statistics Theory and Methods*, **30**, 69–74.
- Ahmadi, J. and Balakrishnan, N. (2012). Outer and inner prediction intervals for order statistics intervals based on current records, *Statistical Papers*, **53**, 789–802.
- Ahmadi, J., MirMostafaei, S. M. T. K. and Balakrishnan, N. (2010). Nonparametric prediction intervals for future record intervals based on order statistics, *Statistics and Probability Letters*, **80**, 1663–1672.
- Arnold, B. C., Balakrishnan, N. and Nagaraja, H. N. (1998). *Records*, Wiley, New York.
- Arnold, B. C., Balakrishnan, N. and Nagaraja, H. N. (2008). *A First Course in Order Statistics (Classic Edition)*, SIAM, Philadelphia.
- Balakrishnan, N., Kamps, U. and Kateri, M. (2009). Minimal repair under a step stress test, *Statistics and Probability Letters*, **79**, 1548–1558.
- Bartoszewicz, J. (2009). On a representation of weighted distributions, *Statistics and Probability Letters*, **79**, 1690–1694.
- Cover, T. M. and Thomas, J. A. (2006). *Elements of Information Theory*, John Wiley and Sons, Inc.
- Csiszár, I. (1963). Eine informationstheoretische Ungleichung und ihre Anwendung auf den Beweis der Ergodizität von Markoffschen Ketten. *Publication of the Mathematical Institute of Mathematics, Hungarian Academy of Sciences*, **8**, 85–108.
- David, H. A. and Nagaraja, H. N. (2003). *Order Statistics*, Third ed., Wiley, Hoboken, NJ.
- Ebrahimi, N., Soofi, E. S. and Zahedi, H. (2004). Information properties of order statistics and spacings, *IEEE Transactions on Information Theory*, **50**, 177–183.
- Fisher, R. A. (1934). The effect of methods of ascertainment upon the estimation of frequencies, *Annals of Eugenics*, **6**, 13–25.
- Gupta, R. C. (1984). Relationship between order statistics and record values and some characterization result, *Journal of Applied Probability*, **21**, 425–430.
- Gupta, R. C. and Ahsanullah, M. (2004). Some characterization results based on the conditional expectation of a function of non-adjacent order statistic (record value), *Annals of the Institute of Statistical Mathematics*, **56**, 721–732.
- Gupta, R. C. and Keating, J. P. (1986). Relations for the reliability measures under length biased sampling, *Scandinavian Journal of Statistics*, **13**, 49–56.
- Jones, M. C. (2004). Families of distributions arising from distributions of order statistics (with discussion), *Test*, **13**, 1–43.
- Kullback, S. and Leibler, R. A. (1951). On information and sufficiency, *Annals of Mathematical Statistics*, **22**, 79–86.
- MirMostafaei, S. T. K. and Ahmadi, J. (2012). General families of distributions arising from distributions of record statistics, Submitted.

- Nevzorov, V. B. (2000). On a limit relation between order statistics and records, *Journal of Mathematical Sciences*, **99**, 1149–1153.
- Patil, G. P. and Rao, C. R. (1977). Weighted distributions: A survey of their application. In P. R. Krishnaiah (Ed.), *Applications of Statistics*, (pp. 383–405). North Holland Publishing Company.
- Patil, G. P. and Rao, C. R. (1978). Weighted distributions and size-biased sampling with applications to wildlife populations and human families, *Biometrics*, **34**, 179–180.
- Rao, C. R. (1965). Weighted distributions arising out of methods of ascertainment. In Classical and Contagious Discrete Distributions, G. P. Patil (Eds). Calcutta: Pergamon Press and Statistical Publishing Society, 320–332.
- Rao, C. R. (1985). Weighted distributions arising out of methods of ascertainment: What population does a sample represent. In the Celebration of Statistics, A. C. Atkinson and S. E. Fienberg eds. Springer Verlag, New York, Chapter **24**, 543–569.
- Shaked, M. and Shanthikumar, J. G. (1994). *Stochastic Orders and Their Applications*, Academic Press, New York.
- Shannon, C. E. (1948). A mathematical theory of communication, *Bell System Technical Journal*, **27**, 379–423 and 623–656.
- Wilks, S. S. (1948). Order statistics, *Bulletin of the American Mathematical Society*, **54**, 6–50.
- Wong, K. M. and Chen, S. (1990). The entropy of ordered sequences and order statistics, *IEEE Transactions on Information Theory*, **36**, 276–284.

Received December 8, 2011; Revised February 21, 2012; Accepted November 1, 2012