Stat Papers (2012) 53:95–106
DOI 10.1007/s00362-010-0316-7

REGULAR ARTICLE

Comparing two sampling schemes based on entropy of record statistics

M. Razmkhah · H. Morabbi · J. Ahmadi

Received: 15 December 2008 / Revised: 21 November 2009 / Published online: 25 February 2010

Abstract In a number of situations such as industrial quality control experiments the only observations are record-breaking data. In this paper, two sampling schemes are used to collect record data: single sample and multisample. The aim of this paper is to investigate which one of them is more efficient in the sense of Shannon information. Several general results are established and it is shown that there is a connection between some reliability properties of the parent distribution and the considered comparison criterion. A number of examples illustrating the results are given.

Keywords Inverse sampling · Hazard rate function · Reversed hazard rate function · Stochastic orders · Shannon information

Mathematics Subject Classification (2000) Primary 62G30 · Secondary 62B10

1 Introduction

Let $\{X_i,\ i \geq 1\}$ be a sequence of independent and identically distributed (iid) continuous random variables. An observation X_j will be called an upper record value if its value is greater than that of all previous observations. Thus X_j is an upper record value if $X_j > X_i$ for all $i < j$. By convention, X_1 is the first upper record value.

M. Razmkhah (✉) · H. Morabbi · J. Ahmadi
Department of Statistics, School of Mathematical Sciences, Ferdowsi University of Mashhad,
P. O. Box 91775-1159, Mashhad, Iran
e-mail: razmkhah_m@um.ac.ir

H. Morabbi
e-mail: h.morabbi@gmail.com

J. Ahmadi
e-mail: ahmadi-j@um.ac.ir

The times at which upper record values appear are given by the random variables T_j which are called record times and are defined by $T_1 = 1$ with probability 1 and for $j \geq 2$, $T_j = \min\{i : X_i > X_{T_{j-1}}\}$ and $U_j = X_{T_j}$. The waiting time between the i-th and $(i+1)$-th upper record value is called the inter-record time (IRT) and is denoted by $K_i = T_{i+1} - T_i$ $(i = 1, 2, \ldots)$. Also, $\Delta_i = K_i - 1$ is the number of trials following the observation U_i before a new record is obtained. Lower record statistics are analogously defined. The reader is referred to Arnold et al. (1998) for more details.

In this paper, we consider two sampling schemes for generating record data:

(i) The experiments are sequentially done, and the only observations available for analysis are record values and their inter record times and sampling is terminated when the n-th record is observed, which is known single sample. We call this design as plan **A**.

(ii) Suppose m independent sequences are sequentially obtained and the only observations available for analysis are record values and their inter record times and sampling is terminated in the i-th sequence, when the k_i-th record is observed, which is known multisample, where $\sum_{j=1}^{m} k_j = n$. We call this design as plan **B**.

In the contexts of record values, the design **A** is well-known as *inverse sampling scheme*. Most researches on the basis of record values have been done in view of *inverse sampling scheme*, see Arnold et al. (1998) and the references therein. In developing the nonparametric maximum likelihood estimation of the underlying distribution based on record data, Samaniego and Whitaker (1988) extended the single sample results in *inverse sampling scheme* to the multisample case, design **B**. They used design **B** since design **A** does not provide sufficient information to estimate parent distribution nonparametrically. They showed that replicated record sequences provide a data set for which nonparametric estimation of the parent distribution is acceptable from both practical and theoretical perspectives, see also Gulati and Padgett (2003). In fact the data values of plan **A** can provide reliable information only in the left (lower records) or in the right (upper record) tails of the sampling distribution.

The concept of Shannon's information (Shannon 1948) plays a central role in information theory and is sometimes taken as measure of uncertainty or ignorance about the outcome of a random experiment. It is known to be of importance in physics (statistical mechanics), mathematics (probability theory), statistical inference (hypothesis testing), electrical engineering (communication theory), and computer science (algorithmic complexity), see Cover and Thomas (1991) for relationship of information theory to other fields. Also, Verdu (1998) gave a list of selected points of tangency of information theory with other fields. If X is a random variable having a continuous pdf f, then the basic uncertainty measure of X is defined as

$$H_X = -\int_{-\infty}^{\infty} f(x) \log f(x) dx. \tag{1}$$

In the literature, H_X is commonly referred to as the *entropy* of X or *Shannon information* measure and used as a tool to determine the amount of information associated with X regarding its parent distribution. Recently, several authors have studied the subject of entropy properties of record statistics. Raqab and Awad (2000) characterized a class of Pareto distributions based on the entropy of record statistics, Zahedi and Shakil (2006) studied the properties of entropies of record values for some common distribution, Baratpour et al. (2007a) investigated the entropy properties of record statistics and showed that the mutual information between record values is distribution free, Baratpour et al. (2007b) presented some characterizations based on entropy of order statistics and record values, Ahmadi and Fashandi (2008) examined the entropy properties of record statistics, especially the difference between entropy of upper and lower bounds of record coverage and obtained several upper and lower bounds. Habibi et al. (2007) investigated testing exponentially based on entropy of record values. No previous work has been done on comparison study of two sampling designs on the basis of entropy. These motivate us to investigate the entropies of record data from two mentioned plans **A** and **B**. We present a comparison study to find out which scheme, **A** or **B**, is more informative, in the sense of Shannon information. Toward this end, we suppose $k_j = k$ $(j = 1, \ldots, m)$, and so we compare the entropy of the first mk record statistics in two sampling schemes **A** and **B**. For convenience of notations, we use $UBGA$, $UBLA$ or $UBEA$ for a distribution F, when entropy of upper records in plan B is greater than, less than or equal to that in plan A, respectively. Also, $UTBGA$, $UTBLA$ or $UTBEA$ is applied for a distribution, if entropy of upper records and their inter record times in plan B is greater than, less than or equal to that in plan A, respectively. Similarly, $LBGA$, $LBLA$, $LBEA$, $LTBGA$, $LTBLA$ and $LTBEA$ are used when the data values are lower record statistics.

In Sect. 2, we obtain the entropy of upper (lower) records and their inter record times. Some general results are derived regarding the entropy of the records and their IRTs of sampling schemes **A** and **B**. Section 3 contains similar results of Sect. 2 for records (upper or lower) alone. In Sect. 4, we consider several common statistical distributions and illustrate the proposed procedure.

2 Results in terms of records and their IRT's

In the following theorem, we obtain an expression for entropy of the first k upper records and their inter record times in terms of reversed hazard rate function.

Theorem 1 *Let $\{X_i,\ i \geq 1\}$ be a sequence of iid continuous random variables with cdf $F(x)$, pdf $f(x)$ and reversed hazard rate function $r(x) = \dfrac{f(x)}{F(x)}$. The joint entropy of the first k upper records and their IRTs is*

$$H_{\mathbf{U},\boldsymbol{\Delta}}(k) = k - \sum_{i=1}^{k} E\left\{\log\left(r\left(F^{-1}(1-e^{-Y_i})\right)\right)\right\}, \tag{2}$$

where $(\mathbf{U}, \boldsymbol{\Delta}) = (U_1, \Delta_1, \ldots, U_{k-1}, \Delta_{k-1}, U_k)$ and Y_i has a gamma distribution with parameters i and 1, i.e., $Y_i \sim \Gamma(i, 1)$.

Proof By the Markov chain property of record statistics, we have

$$H_{\mathbf{U},\mathbf{\Delta}}(k) = H_{U_1} + H_{U_2,\Delta_1|U_1} + \cdots + H_{U_k,\Delta_{k-1}|U_{k-1}}. \tag{3}$$

By (1), the entropy of the first upper record statistic is

$$H_{U_1} = 1 - \int_{-\infty}^{\infty} f(x)\log r(x)dx. \tag{4}$$

Notice that the joint density of the first k upper records and their inter record times is (see Arnold et al. 1998)

$$f_{\mathbf{U},\mathbf{\Delta}}(\mathbf{u},\delta) = \prod_{i=1}^{k} f(u_i)[F(u_i)]^{\delta_i}, \tag{5}$$

where $\delta_k = 0$. Using (5), the conditional density of U_k and Δ_{k-1} given $U_{k-1} = u_{k-1}$ is as follows

$$f_{U_k,\Delta_{k-1}|U_{k-1}}(u_k,\delta|u_{k-1}) = f(u_k)F^{\delta}(u_{k-1}); \quad u_k > u_{k-1}, \quad \delta = 0, 1, \ldots.$$

Therefore, the conditional entropy of U_k and Δ_{k-1} given $U_{k-1} = u_{k-1}$ is

$$\begin{aligned} H_{U_k,\Delta_{k-1}|U_{k-1}}(u_{k-1}) &= -\sum_{\delta=0}^{\infty}\int_{u_{k-1}}^{\infty} f(x)[F(u_{k-1})]^{\delta}[\log f(x) + \delta\log F(u_{k-1})]dx \\ &= -\frac{F(u_{k-1})}{\bar{F}(u_{k-1})}\log F(u_{k-1}) - \frac{1}{\bar{F}(u_{k-1})}\int_{u_{k-1}}^{\infty} f(x)\log f(x)dx, \end{aligned}$$

where $\bar{F}(x) = 1 - F(x)$. Considering the identity

$$\frac{1}{\bar{F}(y)}\int_{y}^{\infty} f(x)\log f(x)dx + \frac{F(y)\log F(y)}{\bar{F}(y)} = \frac{1}{\bar{F}(y)}\int_{y}^{\infty} f(x)\log r(x)dx - 1,$$

we have

$$\begin{aligned} H_{U_k,\Delta_{k-1}|U_{k-1}} &= E[H_{U_k,\Delta_{k-1}|U_{k-1}}(U_{k-1})] \\ &= \int_{-\infty}^{\infty} H_{U_k,\Delta_{k-1}|U_{k-1}}(y) f_{U_{k-1}}(y)dy \\ &= 1 - \int_{-\infty}^{\infty} f_{U_k}(y)\log r(y)dy, \end{aligned} \tag{6}$$

where $f_{U_k}(y)$ is the marginal pdf of U_k and is given by

$$f_{U_k}(y) = \frac{[-\log(\bar{F}(y))]^{k-1}}{(k-1)!} f(y). \tag{7}$$

By Eq. 7 and substituting (4) and (6) in (3), we get

$$H_{\mathbf{U},\mathbf{\Delta}}(k) = k - \sum_{i=1}^{k} \int_{-\infty}^{\infty} \frac{f(x)[-\log \bar{F}(x)]^{i-1}}{\Gamma(i)} \log r(x) dx.$$

Taking $y = -\log \bar{F}(x)$, the result follows. □

It is well-known that in the literature of reliability theory, a distribution F is said to be $IRFR$, $CRFR$ or $DRFR$, if its reversed hazard rate function is increasing, constant or decreasing, respectively. Comparing of sampling schemes A and B based on entropy of upper records and their IRTs in terms of reversed hazard rate function is considered in the next theorem. First, we recall some notions of stochastic ordering that will be used in finding new results, see Shaked and Shanthikumar (2007) for more details.

Definition 1 Let X and Y be two random variables such that $P(X > t) \leq P(Y > t)$ for all $t \in (-\infty, \infty)$, then X is said to be smaller than Y in the usual stochastic order (denoted by $X \leq_{st} Y$).

Definition 2 Let X and Y be two random variables with pdfs f and g, respectively, such that $\frac{g(t)}{f(t)}$ increases in t over the union of the supports of X and Y, then X is said to be smaller than Y in likelihood ratio order (denoted by $X \leq_{lr} Y$).

Lemma 1 (Shaked and Shanthikumar 2007) *For two random variables, X and Y, the following statements are hold:*

(i) If $X \leq_{lr} Y$, then $X \leq_{st} Y$.
(ii) If $X \leq_{st} Y$ and g is an increasing [a decreasing] function, then $g(X) \leq_{st} [\geq_{st}] g(Y)$.
(iii) $X \leq_{st} Y$ if and only if for all increasing function ξ, $E[\xi(X)] \leq E[\xi(Y)]$ for which the expectations exist.

Lemma 2 *Let X and Y be two gamma random variables with shape parameters α_1 and α_2, and scale parameters λ_1 and λ_2, respectively. If $\alpha_1 \leq \alpha_2$ and $\lambda_1 \leq \lambda_2$, then $X \leq_{lr} Y$.*

Theorem 2 *Under the assumptions of Theorem 1, a distribution F is $UTBGA$, $UTBEA$ or $UTBLA$, if it is $IRFR$, $CRFR$ or $DRFR$, respectively.*

Proof By (2), the difference entropy of the first mk upper records and their IRTs of sampling schemes A and B is

$$H^B_{\mathbf{U},\mathbf{\Delta}}(mk) - H^A_{\mathbf{U},\mathbf{\Delta}}(mk) = \sum_{i=1}^{mk} \varphi(i) - m \sum_{i=1}^{k} \varphi(i) = \sum_{i=1}^{m} \sum_{j=1}^{k} \{\varphi\left(j + k(i-1)\right) - \varphi(j)\} \quad (8)$$

for which

$$\varphi(i) = E\left[\log\left(r\left(F^{-1}(1 - e^{-Y_i})\right)\right)\right],$$

where $Y_i \sim \Gamma(i, 1)$. Therefore, by Lemmas 1 and 2, $F^{-1}(1 - e^{-Y_i}) <_{st} F^{-1}(1 - e^{-Y_{i+1}})$. Since $\log x$ is an increasing function, $\varphi(i)$ is an increasing, constant or decreasing function in i, when $r(x)$ is an increasing, constant or decreasing function, respectively. Hence by (8), the result follows. □

Lemma 3 *Under the assumptions of Theorem 1, the joint entropy of the first k lower records and their inter record times is*

$$H_{\mathbf{L},\mathbf{\Delta}'}(k) = k - \sum_{i=1}^{k} E\left\{\log\left(h\left(F^{-1}(e^{-Y_i})\right)\right)\right\}, \quad (9)$$

where $(\mathbf{L}, \mathbf{\Delta}') = (L_1, \Delta'_1, \ldots, L_{k-1}, \Delta'_{k-1}, L_k)$, $h(x) = \frac{f(x)}{\bar{F}(x)}$ *is the hazard rate function of* X *and* $Y_i \sim \Gamma(i, 1)$.

Proof Using the identity

$$\frac{1}{F(y)} \int_{-\infty}^{y} f(x) \log h(x) dx - 1 = \frac{1}{F(y)} \int_{-\infty}^{y} f(x) \log f(x) dx + \frac{\bar{F}(y)}{F(y)} \log \bar{F}(y)$$

the proof is similar to that of Theorem 1. □

In the following corollary some similar results of Theorem 2 in terms of hazard rate function are presented, when the interested statistics are lower records and their inter record times. It is well-known that a distribution F is said to be IFR, CFR or DFR, if its hazard rate function is increasing, constant or decreasing, respectively.

Corollary 1 *Under the assumptions of Theorem 1 and using Lemma 1, a distribution F is LTBLA, LTBEA or LTBGA, if it is IFR, CFR or DFR, respectively.*

3 Results in terms of records without their IRT's

In this section, we consider record values (upper and lower) alone and study the problems presented in the previous section.

Lemma 4 *Under the assumptions of Theorem 1, the joint entropy of the first k upper records is given by*

$$H_U(k) = k - \sum_{i=1}^{k} E\left\{\log\left(h\left(F^{-1}(1-e^{-Y_i})\right)\right)\right\}, \tag{10}$$

where $h(x)$ is the hazard rate function of X and $Y_i \sim \Gamma(i, 1)$.

Proof Habibi et al. (2007), showed that

$$H_{\mathrm{U}}(k) = k - \frac{k(k+1)}{2} - \sum_{i=1}^{k} \int_{-\infty}^{\infty} \frac{f(x)[-\log \bar{F}(x)]^{i-1}}{\Gamma(i)} \log f(x) dx. \tag{11}$$

Using the identity $\log f(x) = \log h(x) + \log \bar{F}(x)$ and transforming by $y = -\log \bar{F}(x)$, the result follows. □

Corollary 2 *Under the assumptions of Theorem 1 and using Lemma 1, a distribution F is $UBGA$, $UBEA$ or $UBLA$, if it is IFR, CFR or DFR, respectively.*

Lemma 5 *The joint entropy of the first k lower records is given by*

$$H_L(k) = k - \sum_{i=1}^{k} E\left\{\log\left(r\left(F^{-1}(e^{-Y_i})\right)\right)\right\}, \tag{12}$$

where $r(x)$ is the reversed hazard rate function of X and $Y_i \sim \Gamma(i, 1)$.

Corollary 3 *Under the assumptions of Theorem 1 and using Lemma 1, a distribution F is $LBLA$, $LBEA$ or $LBGA$, if it is $IRFR$, $CRFR$ or $DRFR$, respectively.*

We would intuitively expect that the entropy of records and their IRT's should be greater than that of those without their IRT's, to investigate this we need the following lemmas.

Lemma 6 *Suppose that $Y_i \sim \Gamma(i, 1)$, then*

$$\sum_{i=1}^{k} E\left\{-\log(1-e^{-Y_i})\right\} = k\left\{1 - \sum_{i=1}^{k-1} \zeta(i+1)\right\} + \sum_{i=1}^{k-1} i\zeta(i+1), \tag{13}$$

where $\zeta(i+1) = \sum_{j=1}^{\infty} \frac{1}{(j+1)^{i+1}}$ is the generalized zeta function.

Proof Notice that

$$\sum_{i=1}^{k} E\left\{-\log(1-e^{-Y_i})\right\} = \sum_{i=1}^{k}\sum_{j=1}^{\infty}\frac{1}{j}E\left\{e^{-jY_i}\right\} = \sum_{j=1}^{\infty}\frac{1}{j^2}\left(1-\frac{1}{(j+1)^k}\right).$$

On the other hand,

$$\begin{aligned}\sum_{j=1}^{\infty}\sum_{i=1}^{k-1}\frac{k-i}{(j+1)^{i+1}} &= k\sum_{j=1}^{\infty}\sum_{i=1}^{k-1}\frac{1}{(j+1)^{i+1}} - \sum_{j=1}^{\infty}\sum_{i=1}^{k-1}\frac{i}{(j+1)^{i+1}} \\ &= \sum_{j=1}^{\infty}\frac{k}{j(j+1)} - \sum_{j=1}^{\infty}\frac{k}{j(j+1)^k} \\ &\quad - \sum_{j=1}^{\infty}\left\{\frac{1}{j^2} - \frac{k}{j^2(j+1)^{k-1}} + \frac{k-1}{j^2(j+1)^k}\right\} \\ &= k - \sum_{j=1}^{\infty}\frac{1}{j^2}\left(1-\frac{1}{(j+1)^k}\right).\end{aligned}$$

Therefore, the result follows. □

Lemma 7 *For $\alpha > 0$, we have*

$$\frac{1}{\alpha 2^{\alpha}} \leq \zeta(\alpha+1) \leq \frac{1}{\alpha}. \tag{14}$$

Proof For $\alpha > 0$ and $n \geq 2$, (see Gut 2007, p. 559), we have

$$\frac{1}{\alpha n^{\alpha}} \leq \sum_{j=n}^{\infty}\frac{1}{j^{\alpha+1}} \leq \frac{1}{\alpha(n-1)^{\alpha}} \leq \frac{2^{\alpha}}{\alpha n^{\alpha}}. \tag{15}$$

Substituting $n = 2$ in (15), the result follows. □

Theorem 3 *Under the assumptions of Theorem 1, for $k \geq 2$, the entropy of the first k upper (lower) records and their IRTs is greater than the entropy of upper (lower) records.*

Proof We show that the result holds for upper records, the case of lower records is similar. By Theorem 1 and Lemma 4, we have

$$H_{\mathbf{U},\mathbf{\Delta}}(k) - H_{\mathbf{U}}(k) = \sum_{i=1}^{k} E\left\{\log\frac{1-e^{-Y_i}}{e^{-Y_i}}\right\},$$

where $Y_i \sim \Gamma(i, 1)$. So, it is sufficient to show that $\sum_{i=1}^{k} E\left\{-\log(1-e^{-Y_i})\right\} < \frac{k(k+1)}{2}$. By the first inequality in (14), we get

$$1 - \sum_{i=1}^{k-1} \zeta(i+1) \leq 1 - \sum_{i=1}^{k-1} \frac{1}{i2^i} < 1 \tag{16}$$

and by the second inequality in (14), we find

$$\sum_{i=1}^{k-1} i\zeta(i+1) \leq k-1. \tag{17}$$

Therefore, by (13), we have

$$\sum_{i=1}^{k} E\left\{-\log(1-e^{-Y_i})\right\} \leq 2k-1 \leq \frac{k(k+1)}{2},$$

which proves the result. □

Remark 1 When the parent distribution is symmetric, the entropy results of this paper based on upper records are similar to those of lower records.

4 Examples

In this section, we present some examples in order to illustrate our results. Let $\{Y_i,\ i \geq 1\}$ be a sequence of iid random variables with an absolutely continuous cdf G and pdf g. Suppose $G(\cdot)$ belongs to the location-scale family of distributions, viz.,

$$G(y) = F\left(\frac{y-\mu}{\sigma}\right) \quad \text{and} \quad g(y) = \frac{1}{\sigma} f\left(\frac{y-\mu}{\sigma}\right), \quad \sigma > 0, \quad \mu \in \mathbb{R},$$

where $F(\cdot)$ and $f(\cdot)$ are the corresponding standard forms (with $\mu = 0$ and $\sigma = 1$). Then, it is evident that $R_n^Y \stackrel{\mathrm{d}}{=} \mu + \sigma R_n^X$, where R_n^Y and R_n^X are the n-th record (upper or lower) value of the Y and X sequences, respectively, and $\stackrel{\mathrm{d}}{=}$ stands for identical in distribution. Then

$$H_{R_n^Y} = H_{R_n^X} + \log \sigma,$$

hence it is obvious that the scale parameter does not any effect on comparing of sampling schemes A and B. So, we can consider the standard form of the corresponding distribution.

Example 1 Let X has a standard normal distribution, then

$$h(x) = \frac{e^{-\frac{x^2}{2}}}{\int_x^\infty e^{\frac{-t^2}{2}} dt}$$

is an increasing function in x. Hence, by Corollaries 1 and 2, normal distribution is $LTBLA$ and $UBGA$, respectively. Using Remark 1, it is also $UTBLA$ and $LBGA$.

Example 2 Let X has an extreme value distribution with cdf

$$F(x) = e^{-e^{-x}}, \quad -\infty < x < \infty.$$

Then the reversed hazard rate function of X is

$$r(x) = \frac{f(x)}{F(x)} = e^{-x},$$

which is decreasing in x. So, by Theorem 2 and Corollary 3, this distribution is $UTBLA$ and $LBGA$, respectively. On the other hand the hazard rate function of X is

$$h(x) = \frac{f(x)}{\bar{F}(x)} = \frac{e^{-x}e^{-e^{-x}}}{1 - e^{-e^{-x}}},$$

where h is an increasing function. Therefore, by Corollaries 1 and 2, extreme value distribution is $LTBLA$ and $UBGA$, respectively.

Example 3 Consider the Weibull distribution with cdf

$$F(x; \alpha) = 1 - e^{-x^\alpha}, \quad x > 0, \ \alpha > 0.$$

Note that the reversed hazard rate function of X is

$$r(x; \alpha) = \frac{f(x; \alpha)}{F(x; \alpha)} = \frac{\alpha x^{\alpha-1} e^{-x^\alpha}}{1 - e^{-x^\alpha}},$$

which is decreasing function in x for positive values of α. So, by Theorem 2 and Corollary 3, the distribution is $UTBLA$ and $LBGA$, respectively.

On the other hand, the hazard rate function of X is given by

$$h(x; \alpha) = \frac{f(x; \alpha)}{\bar{F}(x; \alpha)} = \alpha x^{\alpha-1},$$

where h is an increasing, constant or decreasing function in x, when α is greater than, equal to or less than 1, respectively. Therefore, by Corollary 1 [and Corollary 2], the distribution is $LTBLA$ [and $UBGA$], $LTBEA$ [and $UBEA$] or $LTBGA$ [and $UBLA$], when α is greater than, equal to or less than 1, respectively.

Table 1 Classification of some common distributions based on entropy properties of record data

cdf	Upper records	Lower records	Upper records and their IRT	Lower records and their IRT
$\int_{-\infty}^{x} \frac{1}{\sqrt{2\pi}\sigma} e^{\frac{-(t-\mu)^2}{2\sigma^2}} dt$	*UBGA*	*LBGA*	*UTBLA*	*LTBLA*
$e^{-e^{-\left(\frac{x-\alpha}{\beta}\right)}}$	*UBGA*	*LBGA*	*UTBLA*	*LTBLA*
x^θ				
$\theta \geq 1$	*UBGA*	*LBGA*	*UTBLA*	*LTBLA*
$0 < \theta < 1$	*UBGA**	*LBGA*	*UTBLA*	*LTBLA**
$1 - e^{-\left(\frac{x-\mu}{\sigma}\right)^\alpha}$				
$0 < \alpha < 1$	*UBLA*	*LBGA*	*UTBLA*	*LTBGA*
$\alpha = 1$	*UBEA*	*LBGA*	*UTBLA*	*LTBEA*
$\alpha > 1$	*UBGA*	*LBGA*	*UTBLA*	*LTBLA*
$1 - \left(\frac{\beta}{x}\right)^\alpha$	*UBLA*	*LBGA*	*UTBLA*	*LTBGA*
$1 - (1 + x^\alpha)^{-1}$				
$0 < \alpha \leq 1$	*UBLA*	*LBGA*	*UTBLA*	*LTBGA*
$\alpha > 1$	*UBLA**	*LBGA*	*UTBLA*	*LTBGA**

Notice that, in the special case, $\alpha = 1$, the results of Weibull distribution coincide with those of exponential distribution.

Example 4 A random variable X is said to have a power function distribution if its cdf is

$$F(x; \theta) = x^\theta, \quad 0 < x < 1.$$

Then the reversed hazard rate function of X is

$$r(x; \theta) = \frac{\theta}{x},$$

which is decreasing in x. So by Theorem 2 and Corollary 3, the corresponding distribution is $UTBLA$ and $LBGA$, respectively. On the other hand the hazard rate function of X is

$$h(x; \theta) = \frac{\theta x^{\theta-1}}{1 - x^\theta},$$

which is an increasing function with respect to x for $\theta \geq 1$, whereas for $\theta < 1$ it is not a monotone function. Therefore for $\theta \geq 1$, by Corollaries 1 and 2, the power function distribution is $LTBLA$ and $UBGA$, respectively. Using Eq. 10, we get

$$H^B_{\mathbf{U}}(mk) - H^A_{\mathbf{U}}(mk) = \frac{mk^2(m-1)}{2} + \frac{\theta - 1}{\theta} \sum_{i=1}^{m} \sum_{j=1}^{k} \{\xi\,(j+k(i-1)) - \xi(j)\}\,, \quad (18)$$

where $\xi(i) = E[\log(1-e^{-Y_i})]$ and $Y_i \sim \Gamma(i, 1)$. Notice that the sign of the expression (18) cannot be mathematically determined when $\theta < 1$; Numerical computations in this case indicate that the distribution is $UBGA$. Similarly, for lower records and their IRTs, by Eq. 9, we find

$$H^A_{\mathbf{L},\Delta'}(mk) - H^B_{\mathbf{L},\Delta'}(mk) = \frac{\theta - 1}{\theta} \frac{mk^2(m-1)}{2} + \sum_{i=1}^{m} \sum_{j=1}^{k} \{\xi\,(j + k(i-1)) - \xi(j)\}\,. \quad (19)$$

Using (19), it can be numerically shown that for $\theta < 1$ the distribution is *LTBLA*.

We have considered several other common life distributions and the results are summarized in Table 1. In this table, * means that the result there has been obtained using numerical computations, but has not been proved mathematically.

Acknowledgments The authors would like to thank the referees for their careful reading and useful comments that improved presentation of the manuscript. This paper was partially supported by the Ordered and Spatial Data Center of Excellence of Ferdowsi University of Mashhad.

References

Ahmadi J, Fashandi M (2008) Shannon information properties of the endpoints of record coverage. Commun Stat Theory Methods 37:481–493
Arnold BC, Balakrishnan N, Nagaraja HN (1998) Records. Wiley, New York
Baratpour S, Ahmadi J, Arghami NR (2007a) Entropy properties of record statistics. Statist Pap 48:197–213
Baratpour S, Ahmadi J, Arghami NR (2007b) Some characterizations based on entropy of order statistics and record values. Commun Stat Theory Methods 36:47–57
Cover TM, Thomas JA (1991) Elements of information theory. Wiley-Inter science Publication, New York
Gulati S, Padgett WJ (2003) Parametric and nonparametric inference from record-breaking data. Lecture Notes in Statistics 172. Springer, New York
Gut A (2007) Probability: a graduate course. Springer, New York
Habibi A, Yousefzadeh F, Amini M, Arghami NR (2007) Testing exponentially based on record values. JIRSS 6(1):77–87
Raqab MZ, Awad AM (2000) Characterizations of the Pareto and related distributions. Metrika 52:63–67
Samaniego FJ, Whitaker LR (1988) On estimating population characteristics from record-breaking observations. II: nonparametric results. Naval Res Logist Q 35:221–236
Shaked M, Shanthikumar JG (2007) Stochastic orders. Springer, New York
Shannon CE (1948) A mathematical theory of communication. Bell Syst Tech J 27:379–423
Verdu S (1998) Fifty years of Shannon theory. IEEE Trans Inf Theory 44:2057–2078
Zahedi H, Shakil M (2006) Properties of entropies of record values in reliability and life testing context. Commun Stat Theory Methods 35:997–1010