Journal of the Korean Statistical Society 43 (2014) 263–273

Contents lists available at ScienceDirect

Journal of the Korean Statistical Society

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

Current records and record range with some applications

H.M. Barakat [a,*], E.M. Nigm [a], R.A. Aldallal [b]

[a] Department of Mathematics, Faculty of Science, Zagazig University, Zagazig, Egypt

[b] Department of Mathematics and Statistics, MSA University, Cairo, Egypt

ARTICLE INFO

Article history:
Received 8 January 2013
Accepted 14 September 2013
Available online 8 October 2013

AMS 2000 subject classifications:
primary 62G30
secondary 62G15

Keywords:
Current records
Record range
Stopping time
Recurrence relations
Weak convergence
Weibull distribution
Positive Weibull distribution
Pareto distribution
Negative Pareto distribution

ABSTRACT

In a sequence of independent and identically distributed (iid) random variables, the upper (lower) current records and record range are studied. We derive general recurrence relations between the single and product moments for the upper and lower current records based on Weibull and positive Weibull distributions, as well as Pareto and negative Pareto distributions, respectively. Moreover, some asymptotic results for general current records are established. In addition, a recurrence relation and an explicit formula for the moments of record range based on the exponential distribution are given. Finally, numerical examples are presented to illustrate and corroborate theoretical results.

1. Introduction

Let $\{X_j\}$ be a sequence of independent and identically random variables (rv's) each distributed according to an absolutely continuous cumulative distribution function (cdf) $F_X(x) = P(X \leq x)$ and a probability density function (pdf) $f(x)$. An observation X_j will be called an upper record value if $X_j > X_i$ for every $i < j$. An analogous definition deals with lower record values. In a number of situations, one may store the largest and the smallest X values observed at the times when a new record of either kind (upper or lower) occurs. In this case these records are called current records. Specifically, we call U_n^c and L_n^c the nth upper and lower current records of the sequence X_n, respectively, when the nth record of any kind (either an upper or lower) is observed. It can be noticed that $U_{n+1}^c = U_n^c$ if $L_{n+1}^c < L_n^c$ and that $L_{n+1}^c = L_n^c$ if $U_{n+1}^c > U_n^c$, for all $n = 1, 2, \ldots,$ where by definition, $L_0^c = U_0^c = X_1$. The record range is then defined by $R_n^c = U_n^c - L_n^c$. The record range may also be defined as the nth record range in the sequence of the usual sample range $R_n = \max(X_1, X_2, \ldots, X_n) - \min(X_1, X_2, \ldots, X_n)$, where by definition $R_0^c = 0$ and $R_1^c = R_2$. Note that a new record range is attained once a new upper or lower record is observed.

Record values, as well as current records and record ranges, come up in several life situations. One application is industrial stress and life testing where measurements are taken sequentially and only observations that exceed, or only those that fall below, the current extreme value are recorded. It is interesting to note that there are situations in which only records are observed. Moreover, there are some situations wherein upper and lower records are observed together such as in the case of weather data. In these cases, it would therefore be interesting to consider current records. The current records can be used

* Corresponding author. Tel.: +20 127476652; fax: +20 55 2308213.
E-mail address: hmbarakat@hotmail.com (H.M. Barakat).

1226-3192/$ – see front matter © 2013 The Korean Statistical Society. Published by Elsevier B.V. All rights reserved.
http://dx.doi.org/10.1016/j.jkss.2013.09.004

in general sequential method for model choice and outlier detection involving the record range (Basak, 2000). Actually, both current record values and record range can be detected in several real life situations. For example, the consistency of the production process is required to meet a product's specifications. If the current record range is large, then it is likely that large number of products will lie outside the specifications of the product. Ahmadi and Balakrishnan (2004) obtained distribution-free confidence intervals for quantiles based on current records and record range. Ahmadi and Balakrishnan (2005) established some reliability relationships between the original variable and the corresponding current records. Moreover, Ahmadi and Balakrishnan (2011) considered the case of records and order statistics jointly and discussed the construction of prediction intervals for order statistics. They observed that in the process of obtaining the ordinary record values, one usually observes the current records, and so it is worthwhile to use them in the construction of prediction intervals for order statistics. For more details about current records, record range and their applications, one may refer to Ahmadi and Balakrishnan (2008), Ahmadi, Razmkhaha, and Balakrishnan (2009) and Raqab (2009).

Houchens (1984) used an inductive argument to derive the pdf of U_n^c, L_n^c and R_n^c, based on an arbitrary cdf F_X, (in the sequel we write $U_n^c \parallel X$, $L_n^c \parallel X$ and $R_n^c \parallel X$ to indicate that these statistics are based on the cdf F_X), respectively by

$$f_{U_n^c\|X}(x) = 2^n f_X(x)\left\{1 - \bar{F}_X(x)\sum_{k=0}^{n-1}\frac{[-\log\bar{F}_X(x)]^k}{k!}\right\}, \tag{1.1}$$

$$f_{L_n^c\|X}(x) = 2^n f_X(x)\left\{1 - F_X(x)\sum_{k=0}^{n-1}\frac{[-\log F_X(x)]^k}{k!}\right\} \tag{1.2}$$

and

$$f_{R_n^c\|X}(r) = \frac{2^n}{(n-1)!}\int_{-\infty}^{\infty} f_X(r+x)f_X(x)\{-\log[1-F_X(r+x)+F_X(x)]\}^{n-1}dx,\ 0<r<\infty, \tag{1.3}$$

where $\bar{F}_X(x) = 1 - F_X(x)$.

Remark 1.1. It is worth mentioning that, the cdf of the nth current upper (lower) record is the mixture of the cdf of the $(n+1)$th current upper (lower) record and the cdf $\Gamma_{n+1}(-2\log\bar{F}(x))$ $(1-\Gamma_{n+1}(-2\log F(x)))$, where $\Gamma_n(\theta) = \frac{1}{\Gamma(n)}\int_0^\theta t^{n-1}e^{-1}dt$ is the incomplete gamma function and the mixture constant is 1/2. This fact can be easily derived from the following two relations, which in turn can be easily proved from (1.1) and (1.2), respectively, by routine calculations.

$$F_{U_{n+1}^c\|X}(x) = 2F_{U_n^c\|X}(x) - \Gamma_{n+1}(-2\log\bar{F}(x))$$

and

$$F_{L_{n+1}^c\|X}(x) = 2F_{L_n^c\|X}(x) - (1-\Gamma_{n+1}(-2\log\bar{F}(x))).$$

Since, we have

$$\begin{aligned} f_{-U_n^c\|X}(x) &= f_{U_n^c\|X}(-x) = 2^n f_X(-x)\left[1-\bar{F}_X(-x)\sum_{k=0}^{n-1}\frac{(-\log\bar{F}_X(-x))^k}{k!}\right] \\ &= 2^n f_{-X}(x)\left[1-\bar{F}_{-X}(x)\sum_{k=0}^{n-1}\frac{(-\log\bar{F}_{-X}(x))^k}{k!}\right], \end{aligned}$$

then, we get

$$-U_n^c \parallel X \stackrel{d}{=} L_n^c \parallel -X, \tag{1.4}$$

where "$X \stackrel{d}{=} Y$" means that the rv's X and Y have the same cdf's. The relation (1.4) yields that $U_n^c \parallel X \stackrel{d}{=} -L_n^c \parallel X$, if X is symmetric, i.e. $f_X(x) = f_X(-x) = f_{-X}(x)$.

Houchens (1984) deduced an outstanding representation of $U_n^c \parallel X$, when X has a negative exponential with parameter 2, i.e., $X \sim \text{EX}(2)$. Namely,

$$U_n^c \parallel X \stackrel{d}{=} Y_0 + Y_1 + \cdots + Y_n, \tag{1.5}$$

where Y_i's are independent rv's such that $Y_0 \sim \text{EX}(2)$ and the remaining $Y_i \sim \text{EX}(1)$. An analogous representation for the lower current record can be easily obtained by applying (1.4). Namely, from (1.5), we have

$$-U_n^c \parallel X \stackrel{d}{=} -Y_0 - Y_1 - \cdots - Y_n \stackrel{d}{=} Z_0 + Z_1 + \cdots + Z_n,$$