Characterization of exponential distribution via regression of one record value on two non-adjacent record values

George P. Yanev

Received: 2 February 2009 / Published online: 1 March 2011
© Springer-Verlag 2011

Abstract We characterize the exponential distribution as the only one which satisfies a regression condition. This condition involves the regression function of a fixed record value given two other record values, one of them being previous and the other next to the fixed record value, and none of them are adjacent. In particular, it turns out that the underlying distribution is exponential if and only if given the first and last record values, the expected value of the median in a sample of record values equals the sample midrange.

Keywords Characterization · Exponential distribution · Record values · Median · Midrange

1 Introduction

In 2006, on a seminar at the University of South Florida, Moe Ahsanullah posed the question about characterizations of probability distributions based on regression of a fixed record value with two non-adjacent (at least two spacings away) record values as covariates. We address this problem here.

To formulate and discuss our results we need to introduce some notation as follows. Let $X_1, X_2, \ldots$ be independent copies of a random variable $X$ with absolutely continuous distribution function $F(x)$. An observation in a discrete time series is called a (upper) record value if it exceeds all previous observations, i.e., $X_j$ is a (upper) record value if $X_j > X_i$ for all $i < j$. If we define the sequence $\{T_n, n \geq 1\}$ of
record times by $T_1 = 1$ and $T_n = \min\{j : X_j > X_{T_{n-1}}, j > T_{n-1}\}$, $(n > 1)$, then the corresponding record values are $R_n = X_{T_n}$, $n = 1, 2, \ldots$ (see Nevzorov 2001).

Let $F(x)$ be the exponential distribution function

$$F(x) = 1 - e^{-c(x - l_F)}, \quad (x \geq l_F > -\infty),$$

(1)

where $c > 0$ is an arbitrary constant. Let us mention that (1) with $l_F > 0$ appears, for example, in reliability studies where $l_F$ represents the guarantee time; that is, failure cannot occur before $l_F$ units of time have elapsed (see Barlow and Proschan 1996, p. 13).

We study characterizations of exponential distributions in terms of the regression of one record value with two other record values as covariates, i.e., for $1 \leq k \leq n - 1$ and $r \geq 1$ we examine the regression function

$$E[\psi(R_n) | R_{n-k} = u, R_{n+r} = v], \quad (v > u \geq l_F),$$

where $\psi$ is a function that satisfies certain regularity conditions. Let $\bar{f}_{u,v}$ denote the average value of an integrable function $f(x)$ over the interval from $x = u$ to $x = v$, i.e.,

$$\bar{f}_{u,v} = \frac{1}{v - u} \int_u^v f(t)dt.$$

Yanev et al. (2008) prove, under some assumptions on the function $g$, that if $F$ is exponential then for $1 \leq k \leq n - 1$ and $r \geq 1$,

$$E\left[\frac{g^{(k+r-1)}(R_n)}{k + r - 1} \mid R_{n-k} = u, R_{n+r} = v\right] = \left(\begin{array}{c} k - 1 + r - 1 \\ k - 1 \end{array}\right) \frac{\partial^{k+r-2}}{\partial u^{r-1} \partial v^{k-1}} (\bar{g}_{u,v}),$$

(2)

where $v > u \geq l_F$ and $g'$ is the derivative of $g$. Bairamov et al. (2005) study the particular case of (2) when both covariates are adjacent (one spacing away) to $R_n$. They prove, under some regularity conditions, that if $k = r = 1$, then (2) is also sufficient for $F$ to be exponential. That is, $F$ is exponential if and only if

$$E\left[\frac{g'(R_n)}{} \mid R_{n-1} = u, R_{n+1} = v\right] = \bar{g}_{u,v}, \quad (v > u \geq l_F).$$

Yanev et al. (2008) consider the case when only one of the two covariates is adjacent to $R_n$ and show that, under some regularity assumptions, $F$ is exponential if and only if (2) holds for $2 \leq k \leq n - 1$ and $r = 1$, i.e.,

$$E\left[\frac{g^{(k)}(R_n)}{k} \mid R_{n-k} = u, R_{n+1} = v\right] = \frac{\partial^{k-1}}{\partial v^{k-1}} (\bar{g}'_{u,v}), \quad (v > u \geq l_F).$$