Statistics and Probability Letters 83 (2013) 596–601

Contents lists available at SciVerse ScienceDirect

Statistics and Probability Letters

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

Exponential characterizations motivated by the structure of order statistics in samples of size two

Barry C. Arnold [a], Jose A. Villasenor [b,*]

[a] Department of Statistics, University of California, Riverside, USA

[b] Department of Statistics, Colegio de Postgraduados, Montecillo, MX, Mexico

ARTICLE INFO

Article history:
Received 29 March 2012
Received in revised form 25 October 2012
Accepted 28 October 2012
Available online 7 November 2012

Keywords:
Convolution
Order statistics
Functional equation
Failure rate

ABSTRACT

Motivated by the observation that for a sample of size two from an exponential distribution, the largest order statistic is distributed as a convolution of two independent exponential random variables with distributions differing only in their intensity or rate parameter, a spectrum of related characterizations of the exponential distribution are identified and verified.

1. Introduction

If we consider X_1, X_2, a sample of size two from an exponential distribution, it is well known that the two spacings $X_{1:2}$ and $X_{2:2} - X_{1:2}$ are independent exponentially distributed random variables. Consequently $X_{2:2}$ has as its distribution a convolution of two independent exponential random variables (with differing intensity or rate parameters). For notation, we will write $X \sim \exp(\lambda)$ if the density of X is of the form $f_X(x) = \lambda e^{-\lambda x} I(x > 0)$, and its survival function is given by $\overline{F}_X(x) = P\{X > x\} = e^{-\lambda x},\ x > 0$. In this paper we will refer to λ as an intensity parameter.

Our observation about the structure of the joint distribution of the two spacings corresponding to an exponential sample of size two can be expressed as:

$$X_{2:2} =^d X_1 + \frac{1}{2}X_2$$

where $=^d$ denotes equality in distribution and can be read as "has the same distribution as". Thus $X_{2:2}$ has the same distribution as the convolution of two exponential variables with different intensity parameters. This property holds for exponential samples but may be predicted to be unlikely to hold for samples from other distributions. We will confirm the truth of this assertion, under mild regularity conditions in Section 2. In addition, the standard exponential distribution (corresponding to the case in which $\lambda = 1$) has the, also well known, striking property that $\overline{F}_X(x) = f_X(x),\ x > 0$; i.e., it has a constant failure rate. Combining this distributional property with the convolution property displayed above, we may write the following extensive list of distributional properties that are all satisfied by a sample of size two from a distribution function F with density f and survival function $\overline{F}$, when F is a standard exponential distribution function.

$$X_1 + \frac{1}{2}X_2 =^d \max\{X_1, X_2\}, \tag{1}$$

* Corresponding author.
E-mail address: jvillasr@colpos.mx (J.A. Villasenor).

0167-7152/$ – see front matter
doi:10.1016/j.spl.2012.10.028

$$X_1 + \frac{1}{2}X_2 \quad \text{has density } 2[f(x) - f(2x)], \tag{2}$$

$$X_1 + \frac{1}{2}X_2 \quad \text{has density } 2[\overline{F}(x) - \overline{F}(2x)], \tag{3}$$

$$X_1 + \frac{1}{2}X_2 \quad \text{has density } 2[f(x) - \overline{F}(2x)], \tag{4}$$

$$X_1 + \frac{1}{2}X_2 \quad \text{has density } 2[\overline{F}(x) - f(2x)], \tag{5}$$

$$\max\{X_1, X_2\} \quad \text{has density } 2[f(x) - f(2x)], \tag{6}$$

$$\max\{X_1, X_2\} \quad \text{has density } 2[\overline{F}(x) - \overline{F}(2x)], \tag{7}$$

$$\max\{X_1, X_2\} \quad \text{has density } 2[f(x) - \overline{F}(2x)], \tag{8}$$

$$\max\{X_1, X_2\} \quad \text{has density } 2[\overline{F}(x) - f(2x)], \tag{9}$$

$$f(x) - f(2x) = \overline{F}(x) - \overline{F}(2x). \tag{10}$$

It will be shown that each one of these conditions, on its own, sometimes with mild regularity assumptions on the form of F, is sufficient to guarantee that

$$f(x) = \lambda e^{-\lambda x} I(x > 0) \quad \text{for some } \lambda > 0.$$

We will throughout assume that we are dealing with absolutely continuous positive random variables (thus $F(0) = 0$) with density function $f(x)$.

A convenient survey of other exponential characterizations may be found in Chapter 19 of Johnson et al. (1994). See also Arnold and Huang (1995).

Remark 1. Characterizations are particularly of interest when they shed light on the consequences of certain distributional assumptions and/or can be used to assess the plausibility of such assumptions via suitable tests of hypotheses. For example, consider the characterization based on Eq. (1). Eq. (1) will hold if a parallel system of two identical components exhibits the same reliability as a single component provided with a cold standby component with doubled failure rate. If such a situation is deemed to be plausible, then the assumption of an exponential distribution for the failure times will lead to an acceptable model. On the other hand, if a sample of n X's is available then one can randomly divide the data set into four subsets, relabeled as

$$U_1, U_2, \ldots, U_{n/4}, \qquad V_1, V_2, \ldots, V_{n/4}, \qquad W_1, W_2, \ldots, W_{n/4}, \qquad Z_1, Z_2, \ldots, Z_{n/4}.$$

Then for $i = 1, 2, \ldots, n/4$ define $S_i = \max\{U_i, V_i\}$ and $T_i = W_i + (1/2)Z_i$. The S's and the T's will have a common distribution if and only if the original X's have an exponential distribution. Any standard two-sample non-parametric test can be used to compare the sample distribution functions of the S's and the T's, to provide evidence regarding the acceptability of the exponential model.

2. The characterizations

In all theorems below, it is assumed that X_1, X_2 is a sample of size two from a distribution F, assumed to be absolutely continuous with $F(0) = 0$, with density function $f(x)$, and with Laplace transform $\zeta(t) = E(e^{-tX_1})$. We begin by recalling a useful Lemma which will be used in two of the theorems.

Lemma 1. *If a function $g : [0, \infty] \longrightarrow (-\infty, \infty)$ has a right derivative at 0 denoted by $g'(0)$ and satisfies*

$$g(t) = 2g(t/2), \quad \textit{for every } t \geq 0,$$

then $g(t) = tg'(0)$ for every $t \geq 0$.

Proof. For any $t > 0$, since $g(t) = 2g(t/2)$ it follows by induction that

$$g(t) = 2^k g(t/2^k) \quad \text{for all } k = 1, 2, \ldots.$$

Consequently

$$g(t) = \lim_{k\to\infty} 2^k g(t/2^k) = \lim_{k\to\infty} t\frac{g(t/2^k)}{t/2^k} = tg'(0). \quad \square$$

Rather than present ten separate theorems, one for each of the candidate characterization conditions (1)–(10), we will group those which use the same regularity condition.