Applied Mathematical Modelling 40 (2016) 6775–6784

Contents lists available at ScienceDirect

Applied Mathematical Modelling

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

An exponentiated geometric distribution

S. Nadarajah [a,*], S.A.A. Bakar [b]

[a] School of Mathematics, University of Manchester, Manchester M13 9PL, UK
[b] Institute of Mathematical Sciences, University of Malaya, 50603 Kuala Lumpur, Malaysia

ARTICLE INFO

Article history:
Received 27 April 2013
Revised 3 November 2015
Accepted 9 November 2015
Available online 29 November 2015

Keywords:
Estimation
Exponentiated geometric distribution
Moments

ABSTRACT

The exponentiated exponential distribution is one of the most seminal distributions of the 20th century. Here, we propose a discrete exponentiated exponential distribution. We derive its mathematical properties and procedures for estimation by common methods. The estimation procedures are assessed by simulation. Finally, the proposed distribution is compared with one of the most recent discrete distributions using two real data sets on insurance.

1. Introduction

Discrete distributions arise naturally in insurance as models for claim counts. However, traditional discrete distributions like the geometric, negative binomial and Poisson distributions are not flexible enough. Some new discrete distributions have been proposed recently as flexible models in insurance. These new distributions are often taken to be discrete versions of well known continuous distributions.

One of the most seminal continuous distributions introduced in the 20th century is the exponentiated exponential distribution [1]. The exponentiated exponential distribution is most seminal in that it has inspired the development of many other distributions, including other exponentiated type distributions, beta exponentiated type distributions due to Eugene et al. [2], gamma type distributions due to Zografos and Balakrishnan [3], Kumaraswamy type distributions due to Cordeiro and Castro [4], beta extended type distributions due to Cordeiro et al. [5], gamma type distributions due to Ristić and Balakrishnan [6], exponentiated Kumaraswamy type distributions due to Lemonte et al. [7], and so on.

The exponentiated exponential distribution has also received widespread applications. Its applications have included: models to determine bout criteria for analysis of animal behavior, design rainfall estimation in the Coast of Chiapas, analysis of Los Angeles rainfall data, software reliability growth models for vital quality metrics, models for episode peak and duration for eco-hydro-climatic applications, estimating mean life of power system equipment with limited end-of-life failure data, and cure rate modeling. For a comprehensive account of mathematical properties and applications of the exponentiated exponential distribution, we refer the readers to Nadarajah [8].

Let $N_1, N_2, \ldots, N_\alpha$ denote independent claim counts of α different insurance companies. Assume that each N_i is a geometric random variable with parameter θ. Let $X = \max(N_1, N_2, \ldots, N_\alpha)$ denote the maximum claim amount. The cumulative distribution function of X is

* Corresponding author. Tel.: +4401614732941.
E-mail addresses: mbbsssn2@manchester.ac.uk (S. Nadarajah), saab@um.edu.my (S.A.A. Bakar).

http://dx.doi.org/10.1016/j.apm.2015.11.010

$$\begin{aligned} F(x) &= P(\max(N_1, N_2, \ldots, N_\alpha) \le x) = P(N_1 \le x, N_2 \le x, \ldots, N_\alpha \le x) \\ &= P(N_1 \le x)P(N_2 \le x) \cdots P(N_\alpha \le x) = [P(N_1 \le x)]^\alpha \\ &= \left[1 - (1-\theta)^x\right]^\alpha \end{aligned} \tag{1}$$

for $0 < \theta < 1$, $\alpha > 0$ and $x = 0, 1, 2, \ldots$. We shall refer to X as an exponentiated geometric random variable and its distribution as the exponentiated geometric distribution.

The corresponding probability mass and survival functions are

$$p(x) = F(x) - F(x-1) = \left[1 - (1-\theta)^x\right]^\alpha - \left[1 - (1-\theta)^{x-1}\right]^\alpha \tag{2}$$

and

$$S(x) = 1 - F(x) = 1 - \left[1 - (1-\theta)^x\right]^\alpha, \tag{3}$$

respectively, for $x = 1, 2, \ldots$ and for $x = 0, 1, 2, \ldots$, respectively. The corresponding failure rate function is

$$h(x) = \frac{p(x)}{S(x)} = \frac{\left[1 - (1-\theta)^x\right]^\alpha - \left[1 - (1-\theta)^{x-1}\right]^\alpha}{1 - \left[1 - (1-\theta)^x\right]^\alpha} \tag{4}$$

for $x = 1, 2, \ldots$. The probability mass function in (2) is useful for maximum likelihood estimation of the parameters θ and α, see Section 3.2. The survival function in (3) is useful for censored maximum likelihood estimation of the parameters θ and α, see Section 3.4. The failure rate function in (4) assesses the ability of the distribution to model failure times, see Section 2.2.

Note that $p(1) = \theta^\alpha$, $p(\infty) = 0$, $h(1) = \theta^\alpha / (1 - \theta^\alpha)$ and $h(\infty) = \theta / (1 - \theta)$. A particular case of (1) for $\alpha = 1$ is the geometric distribution.

Having closed form expressions for its cumulative distribution and failure rate functions is an attractive feature of the exponentiated geometric distribution. Poisson, binomial, or the negative binomial distributions do not have closed form expressions for their cumulative distribution and failure rate functions.

The exponentiated geometric distribution appears to be new. However, Jiang ([9], Section 3.2.2) mentions briefly some shape properties of the exponentiated geometric distribution. We are aware of no other papers in the literature, where the exponentiated geometric distribution has been studied in a statistical sense.

The rest of this note is organized as follows. Various mathematical properties of (1) are derived in Section 2. Estimation procedures by two common methods are derived in Section 3. Applications involving insurance data sets are discussed in Section 4.

2. Mathematical properties

The mathematical properties of the exponentiated geometric distribution derived are: expansions (Section 2.1), shape properties (Section 2.2), quantile function (Section 2.3), probability generating function (Section 2.4), moment generating function (Section 2.4), moments (Section 2.5), order statistics (Section 2.6) and distribution of range (Section 2.7).

Some of the given expressions involve infinite series: namely, (6), (7), (9), (10), (13) and (14). Extensive computations not reported here showed that the relative errors between (6), (7), (9), (10), (13) and (14) and their versions with the infinite series in each truncated at 20 did not exceed 10^{-20}. This shows that (6), (7), (9), (10), (13) and (14) can be computed for most practical uses with their infinite sums truncated at 20. The computations were performed using Maple. Maple took only a fraction of a second to compute the truncated versions of (6), (7), (9), (10), (13) and (14). The computational times for the truncated versions were significantly smaller than those for the untruncated versions.

2.1. Expansions

Some of the mathematical properties of (1)–(4) cannot be expressed in closed form. In these cases, it is useful to have expansions for the probability mass and cumulative distribution functions. Using the binomial expansion

$$(1+a)^\alpha = \sum_{j=0}^{\infty} \binom{\alpha}{j} a^j, \tag{5}$$

we can express (1) and (2) as

$$F(x) = \left[1 - (1-\theta)^x\right]^\alpha = \sum_{j=0}^{\infty} \binom{\alpha}{j} (-1)^j (1-\theta)^{xj} \tag{6}$$

and