Journal of Computational and Applied Mathematics 311 (2017) 171–182

Contents lists available at ScienceDirect

Journal of Computational and Applied Mathematics

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

The Weibull–Conway–Maxwell–Poisson distribution to analyze survival data

Ramesh C. Gupta*, Jie Huang

Department of Mathematics and Statistics, University of Maine, Orono, ME 04469-5752, United States

ARTICLE INFO

Article history:
Received 24 January 2016
Received in revised form 20 April 2016

Keywords:
Maximum likelihood estimation
Score test
Exponential-Poisson
Exponential COM–Poisson distribution

ABSTRACT

In this paper, we consider the distribution of life length of a series system with random number of components, say M. Considering the distribution of M as COM–Poisson, a Weibull–COM–Poisson distribution (WCOMP) is developed. The COM–Poisson is a generalization of the Poisson distribution having one extra parameter. The structural properties of the resulting distribution are presented and the maximum likelihood estimation of the parameters is investigated. Extensive simulation studies are carried out to study the performance of the estimates. A score test is developed to test the importance of the extra parameter. For illustration, three real data sets are examined and it is shown that the WCOMP model, presented here, fits better than the exponential Poisson and the exponential-COMP distributions.

1. Introduction

In life testing and survival analysis, sometimes the components are arranged in series or parallel system and the number of components is initially unknown. In the competing risk analysis, the individual times of events are unknown, but we observe $T = \min(Y_1, Y_2, \ldots, Y_M)$ where $Y_1, Y_2, \ldots, Y_M$ are independent, identically distributed random variables and M is considered as random. The distribution of the life length of series and parallel system with random number of components has been studied by various authors. For example, suppose for a patient, M denotes the number of carcinogenic cells (often called clonogens) left active after the first treatment and let M have a discrete distribution having mass at $\{1, 2, 3, \ldots\}$. Let $Y_i, i = 1, 2, \ldots, M$ denote the incubation time for the ith clonogenic cell. For a given $M \geq 1$, we assume that one (all) out of M latent factors need to activate for the subjects to fail. Then the time to fail is $T = \min(Y_1, Y_2, \ldots, Y_M)$ or $T = \max(Y_1, Y_2, \ldots, Y_M)$. Cooner et al. [1] call such system as series (parallel) system, the first (last) activation scheme. See also Ibrahim et al. [2].

In cure rate models, in addition to the individuals who are uncured, the interest is in estimating the probability of cure, called the cure rate. In that case M is also allowed to take the value 0. In the last ten years, there has been a considerable interest in studying the cure rate models from different angles. The reader is referred to Borges et al. [3], Rodrigues et al. [4], Rodrigues et al. [5], Fonseca et al. [6], Balakrishnan and Pal [7,8], Peng and Xu [9] and Cancho et al. [10].

In the literature various distributions of M have been considered. Cooner et al. [1], Chen et al. [11], Kus [12] and Karlis [13] have considered the distribution of M as Poisson. Cordeiro et al. [14] and Rodrigues et al. [4] studied the resulting model where M has a Conway–Maxwell–Poisson (COMP) distribution. The COMP distribution is a generalization of the Poisson distribution having one more parameter which allows for underdispersion or overdispersion. Tahmasbi

* Corresponding author.
E-mail address: rcgupta@maine.edu (R.C. Gupta).

http://dx.doi.org/10.1016/j.cam.2016.06.035
0377-0427/

and Rezaei [15] considered the above model having logarithmic series distribution. More generally, Morais and Barreto-Souza [16] considered the distribution of M as power series.

The aim of the paper is to provide an alternative model to analyze survival data. More precisely, in cancer studies, the number of carcinogenic cells left active after the treatment is unknown and hence this number has to be modeled by an appropriate discrete distribution. In most of the studies, Poisson distribution has been used to model this unknown number. The validity of this rationale was questioned by several authors including Tucker et al. [17]. Replacing the Poisson distribution by COMP distribution provides a more flexible model because the COMP distribution takes care of both the overdispersion and the underdispersion. On the other hand, in the case of Poisson distribution, the dispersion parameter is unity, an assumption which is not satisfied in most cases.

Thus, in the present work, we shall consider the distribution of M as COM–Poisson (COMP), which is a generalization of the Poisson distribution having one more parameter. Also, we replace the exponential distribution with a Weibull distribution resulting in a more flexible model having decreasing, increasing and upside bath tub shaped distribution depending on the shape parameter of the Weibull distribution. Our model will be compared to Kus [12] and Cordeiro et al. [14] models in terms of fitting and other characteristics. The probability function of the COMP distribution is given by

$$P(M=j;\lambda,\nu)=\frac{\lambda^j}{Z(\lambda,\nu)(j!)^\nu},\quad j=0,1,\ldots, \tag{1.1}$$

where $Z(\lambda,\nu)=\sum_{j=0}^{\infty}\lambda^j/(j!)^\nu$, $\lambda>0$, $\nu>0$. Note that the probability mass function of the COMP distribution is log-concave and hence it has an increasing failure rate. Thus, our model is a generalization of Kus [12] model in two directions. We replace the baseline exponential distribution by a Weibull distribution and Poisson distribution by a COMP distribution resulting in two more parameters. It is also an extension of Cordeiro et al. [14] model.

The maximum likelihood estimation of the parameters is studied. Rao's score test is developed to test the importance of the additional parameters γ and ν. Three examples are presented to illustrate the procedure and support the model. The paper is organized as follows: In Section 2, we develop the model and derive the survival function. Some structural properties of the model are also studied. In Section 3, the maximum likelihood estimation of the parameters is considered. A score test, for testing the importance of the additional parameters ν and γ, is developed. We illustrate the flexibility of the proposed model in Section 4 using three data sets which demonstrate the importance of the additional parameters. Simulation studies are carried out in Section 5 to investigate the performance of the estimates. Finally, in Section 6, some conclusion and comments are presented.

2. Derivation of the WCOMP model

2.1. The WCOMP distribution

The exponential-Poisson (EP) distribution was introduced by Kus [12], which compounds an exponential distribution with a Poisson distribution. The Conway–Maxwell–Poisson (COMP) distribution, discussed recently by Shmueli et al. [18], Lord et al. [19] and Seller and Shmueli [20], generalizes the Poisson distribution allowing for under-dispersion as well as over-dispersion. The COMP distribution with parameters $\lambda>0$ and $\nu\geq 0$, say $COMP(\lambda,\nu)$, has probability mass function (pmf) given by (1.1).

Cordeiro et al. [14] presented an extension of Kus [12] by compounding an exponential distribution with the COMP distribution and defined a three parameter distribution referred to as the exponential-Conway–Maxwell–Poisson (ECOMP) distribution. We have also replaced the exponential distribution by a Weibull distribution. So, in this paper, we present the Weibull Conway–Maxwell–Poisson (WCOMP) distribution which includes the EP and the ECOMP models as special sub-models.

Let the discrete random variable M in (1.1) be zero truncated with pmf written as

$$P(M=j;\lambda,\nu)=\frac{\lambda^j}{[Z(\lambda,\nu)-1]\,(j!)^\nu},\quad j=0,1,\ldots \tag{2.1}$$

Suppose that $\{Y_i\}_{i=1}^M$ are independent and identically distributed (iid) random variables having the Weibull distribution, say $W(\beta,\gamma)$, with scale parameter $\beta>0$ and shape parameter $\gamma>0$. The density function of $W(\beta,\gamma)$ is given by

$$f_W(x;\beta,\gamma)=\gamma\beta^\gamma x^{\gamma-1}e^{-(\beta x)^\gamma},\quad x>0.$$

We assume that the random variables, $Y_i's$, are independent of M. The random variable $X=\min\{Y_i\}_{i=1}^M$ defines the WCOMP distribution, denoted by $WCOMP(\beta,\lambda,\nu,\gamma)$, for which the conditional density function of X (given $M=j$) is given by $f(x|j;\beta,\gamma)=\beta\gamma j(\beta x)^{\gamma-1}e^{-(\beta x)^\gamma j}$. Thus, the WCOMP distribution can be used to model the minimum lifetimes of Weibull random variables.

The marginal probability density function (pdf) of X becomes

$$f(x;\boldsymbol{\theta})=\frac{\beta\gamma(\beta x)^{\gamma-1}}{Z(\lambda,\nu)-1}\sum_{j=1}^{\infty}\frac{j\left[\lambda e^{-(\beta x)^\gamma}\right]^j}{(j!)^\nu},\quad x>0 \tag{2.2}$$