Available online at www.sciencedirect.com
ScienceDirect
Mathematics and Computers in Simulation 98 (2014) 18–30
MATHEMATICS AND COMPUTERS IN SIMULATION
www.elsevier.com/locate/matcom

Original article

Double-looped maximum likelihood estimation for the parameters of the generalized gamma distribution

Hulya Yilmaz [a], Hakan S. Sazak [b,*]

[a] *Department of Biostatistics, Eskişehir Osmangazi University, 26000 Eskişehir, Turkey*
[b] *Department of Statistics, Ege University, 35100 İzmir, Turkey*

Received 13 August 2012; received in revised form 18 November 2013; accepted 3 December 2013
Available online 12 December 2013

Abstract

The generalized gamma distribution (GGD) is a very popular distribution since it includes many well known distributions. Estimation of the parameters of the GGD is quite problematic because of the complicated structure of its density function. We introduce two new estimation methods called maximum likelihood with goodness of fit test (MLGOFT) and double-looped maximum likelihood (ML) estimation. We show through simulations under several situations that the MLGOFT method is more efficient than the Method of Moments with goodness of fit test (MMGOFT) technique especially for small and moderate sample sizes whereas the double-looped ML is the superior estimation method for all cases. The double-looped ML method is also very fast, practical and straightforward.

Keywords: Generalized gamma distribution; Method of moments; Double-looped maximum likelihood; Parameter estimation; Chi-square goodness of fit test

1. Introduction

The generalized gamma distribution (GGD) was introduced by Stacy [23]. Its probability density function is given by

$$f(x;\ell,a,c)=\frac{|c|\,\exp\{-(x/a)^c\}x^{\ell c-1}}{\Gamma(\ell)a^{\ell c}},\quad x>0,\quad \ell>0,\quad a>0,\quad c\in R$$

where $\Gamma(\cdot)$ is the gamma function, ℓ and c are shape parameters and a is the scale parameter. The GGD is a very flexible distribution since it reduces to well-known distributions like Weibull, gamma, chi-square, exponential, etc. for specific values of the parameters ℓ, a and c. For detailed information on the properties and the special cases of GGD, see Stacy and Mihram [25], Kleiber and Kotz [15] and Khodabin and Ahmadabadi [14]. GGD is used in

* Corresponding author. Tel.: +90 232 3111725; fax: +90 232 3881890.
E-mail addresses: hulyayilmaz@ogu.edu.tr (H. Yilmaz), ssazak@yahoo.com, hakan.savas.sazak@ege.edu.tr (H.S. Sazak).

0378-4754/$36.00
http://dx.doi.org/10.1016/j.matcom.2013.12.001

many areas as economics and actuarial science for modeling various income and loss distributions [15], industrial and medical life testing for modeling life time of various products [5,17], health costs [19] or civil engineering [21]. To be able to use GGD in statistical models, one has to estimate the parameters in its probability density function.

Although the estimation for some specific cases of GGD can be very straightforward, it is not easy to give a solution for GGD since it is a very broad family of continuous univariate probability distributions and there are three unknown parameters, two of which are shape parameters. Many studies have been made for estimating the parameters of the GGD. Since almost all of the proposed methods are quite complex, this topic is still an open area. In general, two estimation techniques, maximum likelihood (ML) and method of moments (MM), were employed. Fisher [6] suggested the use of ML because MM may be inefficient for Pearson type III distributions. Bai et al. [1] and Bowman and Shenton [3] also pointed out the possible inefficiency of MM estimators and suggested the use of ML estimators because of their desirable properties. Stacy and Mihram [25] studied both ML and MM estimation techniques but found that no closed solution is available for both techniques since simultaneous nonlinear equations have to be solved. They suggested a graphical technique which is not very practical. Parr and Webster [20], Harter [10], Hager and Bain [8], Lawless [16], Kleiber and Kotz [15] and Hirose [11] worked on ML estimation. They reported that ML estimation technique is extremely difficult because of the nonlinear terms in the likelihood equations. Hager and Bain [8], Hager et al. [9] and Lawless [16] remarked that solution to the likelihood equations may not be obtained for a sample size less than 400. Harter [10] proposed an iterative procedure for solving these equations. On the other hand, Stacy [24] mentioned that these equations can have multiple roots. In general, iterative methods to solve the nonlinear likelihood equations are very problematic because of nonconvergence, convergence to wrong values, multiple roots etc. See for example Barnett [2], Lee et al. [18], Puthenpura and Sinha [22] and Vaughan [28]. Wingo [29] described an algorithm which uses a heuristic root isolation method developed by Jones et al. [13] to solve the ML equations but it can work only when the shape parameter c is positive. Hirose [11] proposed an estimation technique utilizing model augmentation by the continuation method but his technique requires iteration and it is quite complex. MM can also be very problematic because different sets of parameters conduce to same density function [7,12,15]. Cohen and Whitten [4] proposed a method called the modified moment estimation method which loops on parameter c using some reference charts and tables but their method also requires iterations and involves lengthy calculations. Huang and Hwang [12] worked on the moments of the GGD to estimate the three parameters of this distribution but they provided estimators in closed forms only for two parameters given that the third one is known, in other words, for some special distributions such as gamma and Weibull. Gomes et al. [7] developed a straightforward algorithm which has a loop on parameter c for estimation. This algorithm uses the idea of Cohen and Whitten [4] but also uses the power transformation in order to do estimation in the GGD. Khodabin and Ahmadabadi [14] worked on the estimation of the parameters of the GGD but they focused on estimating the parameters ℓ and a.

The aim of this paper is to derive a simple, straightforward and efficient estimation method for the parameters of the GGD. For this purpose we derive two estimation methods, ML with goodness of fit test (ML GOFT) and double-looped ML estimation. Since the method proposed by Gomes et al. [7] seems to be the only simple and straightforward method for estimating the parameters of GGD so far, we will calculate the relative efficiencies of the estimators produced by the ML GOFT and the double-looped ML methods w.r.t. the estimators produced by their method. In their method, they define an interval for the shape parameter c and then use $Y = X^c$ transformation to obtain gamma distribution. For each c value, they estimate the parameters of gamma distribution by using its own moments, and then calculate and keep the p-value of the chi-square goodness of fit test to observe how well the corresponding transformed Y variable follows gamma(ℓ, a^c) distribution. At the end of the loop, the estimated parameter group (ℓ, a, c), at which the p-value is highest, is the estimated parameter group of the GGD. Gomes et al. [7] conducted simulations for some special parameter values which correspond to well-known distributions such as Weibull, gamma, exponential etc. They gave all the resulting values of the parameters based on 5 simulations for a sample size of 10,000. Let us call their estimation method MM GOFT.

In the rest of our paper, we introduce ML GOFT and double-looped ML estimation methods and investigate their efficiencies. In Section 2, we describe these methods and show how they are implemented. The results of the simulations which are performed to investigate the efficiencies of the introduced methods are given in Section 3. The final section includes some concluding remarks and suggestions.