Mathematical Statistician and Engineering Applications
ISSN: 2094-0343
2326-9865

Asymptotic Normality of the Estimators of the Parameters in Perturbed Gamma Model Using EM Algorithm.

Dr. RaveendraNaika T[1]

Assoc. Professor, Dept. of Statistics, Maharani's Science College for Women, Bengaluru -01
raviisec@yahoo.co.uk

Dr. Shivanna B K[2]

Assoc. Professor, Dept. of Statistics, Maharani's Science College for Women, Mysuru
shivustats@gmail.com

Mr. Kishore Kumar J[3]

Assist. Professor, Dept. of Statistics, Maharani's Science College for Women, Bengaluru -01

kishorestat@gmail.com

Article Info
Page Number: 6964-6971
Publication Issue:
Vol. 71 No. 4 (2022)

Article History
Article Received: 25 March 2022
Revised: 30 April 2022
Accepted: 15 June 2022
Publication: 19 August 2022

Abstract
In this article it has been considered that the model which is the sum of two independent processes are gamma and uniform distribution and the mixture model is known as perturbed gamma model. Based on independent sample of the perturbed gamma model observed at irregular cases and when the missing data arises the expectation and maximization method is applied to estimate the unknown parameters of the model using the maximum likelihood estimator. A simulation study has been carried out to demonstrate the asymptotic normality of the estimators over three-dimensional graph in bivariate normal distribution using R software (maxLik/optim function).
Keywords: Maximum likelihood estimator, EM algorithm, Bivariate normal distribution, Gamma distribution.

1.1 Introduction

The gamma distribution is useful in actuarial modeling. Due to its mathematical properties, there is considerable flexibility in the modeling process. Since it has two parameters (a scale parameter and a shape parameter), the gamma distribution is capable of representing a variety of distribution shapes and dispersion patterns. A social group under study consists of both suitable and unsuitable people for the transplantation of an organ. Then, the lifetime after the transplantation of an individual coming from this group turns out to be a mixture. The maximum likelihood estimation of the parameters involved in the mixture model using the EM algorithm discussed by G.Nanjundan and T. Raveendra Naika (2011) in Maximum Likelihood Estimation of Parameters in a Perturbed Exponential Model. Laurent Bordes, Christian Paroissin & Ali Salami (2016) discussed parametric inference in a perturbed gamma degradation process and estimated the parameters using method of moment estimation. Massimiliano Giorgio , Agostino Mele , Gianpaolo Pulcini (2018) proposes and illustrates a new perturbed gamma degradation process where the measurement error is modeled as a non-Gaussian random variable that depends stochastically on the actual degradation level. The expression of the likelihood function for a generic set of noisy degradation measurements is derived in perturbed gamma degradation process with degradation dependent non-Gaussian measurement errors.

Suppose a company is interested in the duration of the telephone calls going out from its office. Let us assume that a caller does not continue speaking beyond a known time *s*, if the call goes to a wrong person and the duration of such a call is uniformly distributed over $(0, s)$. On the other hand, if call goes to a right person, then the duration of a call is distributed gamma with parameter θ and β (where β is assumed to be known). Further suppose that the proportion of a calls goes to wrong a person from this office is φ.

If a caller continues to speaking beyond time *s*, then definitely the call goes to a right person. On the other hand, if a caller does not continue to speaking beyond time '*s'*, then it cannot be definitely classified as whether the call goes to a right person or a wrong person. In other words, if the duration of telephone call *X* is less than or equal to *s*, then the information about the call is missing whether it goes to a right person or a wrong person.

Then, the duration of telephone calls (*X*) going out from an office of a company has the following probability density function

$$f(x) = \begin{cases} \frac{\varphi}{s} + (1-\varphi)\, \frac{\theta}{\beta}(\theta x)^{\beta-1}\exp(-\theta x), \ 0 < x \le s \\ (1-\varphi)\frac{\theta}{\beta}(\theta x)^{\beta-1}\exp(-\theta x), \ x > s, \ 0<\varphi<1, \ \theta>0, \beta>0 \end{cases}$$

(1.1)

where $f_1(x) = \frac{1}{s}, \ 0 < x \le s$ and $f_2(x) = f_2(x;\theta,\beta) = \frac{\theta}{\beta}(\theta x)^{\beta-1}\exp(-\theta x), \ x > 0, \ (\theta,\beta) > 0.$ is Gamma distribution and θ is a shape/location parameter and β is the rate/scale parameter, β is assumed to be known. Equation (1.1) is known as perturbed Gamma density.

1.2 Maximum Likelihood Estimation

Let $\underline{X} = (X_1, X_2, \ldots, X_n)$ be the random sample from (1.1). Then, the likelihood of θ and φ given the sample $\underline{X}$ can be written as

$$L(\theta,\varphi;\underline{x}) = \prod_{j=1}^{n}\left(\frac{\varphi}{s} + (1-\varphi)\frac{\theta}{\beta}(\theta x_j)^{\beta-1}\exp(-\theta x_j)\right)^{1-y_j}\left((1-\varphi)\frac{\theta}{\beta}(\theta x_j)^{\beta-1}\exp(-\theta x_j)\right)^{y_j}, \qquad (1.2)$$

where $y_j = \begin{cases} 1, if X_j > s \\ 0, if X_j \le s \end{cases}$.

It can be easily seen that the equations $\frac{\partial L}{\partial \theta} = 0$ and $\frac{\partial L}{\partial \varphi} = 0$ do not yield closed form expression for the maximum likelihood estimators (MLEs) of θ and φ. Hence, for a given sample, the maximum likelihood estimates of θ and φ could be computed using one of the numerical iterative procedures like Newton-Raphson, Fletcher-Reeves etc. But these procedures are analytically and computationally tedious in the case of (1.2).

1.3 The EM Algorithm

When the likelihood functions have complicated structures and their maximization by numerical methods is difficult, the MLEs of the parameters can be computed by the Expectation - Maximization (EM) algorithm. It is popular and remarkably simple to implement. It is an iterative procedure and there are two steps in each of the iterations, namely the Expectation Step (E-step) and the Maximization step (M-step). The EM algorithm was developed by Dempster, Laird, and Rubin (1977) who synthesized an earlier formulation in many particular cases and gave a general method of finding the MLEs in a variety of situations. Since then the EM algorithm has been applied to a variety of statistical problems such as resolution of mixtures, multi-way contingency tables, variance component estimation, and factor analysis. It has also found applications in specialized areas like genetics, medical imaging, and neural networks. McLachlan and Krishnan (2008) discuss the EM algorithm and its extensions to a variety of problems in detail while Krishnan (2004) gave a brief introduction to the algorithm with examples.

We now need to rewrite the likelihood so as to accommodate the missing data to apply the EM algorithm to compute the MLEs.

Let $Z_j = \begin{cases} 1, & \text{if the } j-\text{th sampled call goes to a right person} \\ 0, & \text{otherwise.} \end{cases}$

Then, we have $P(Z_j = 1) = 1 - \varphi = 1 - P(Z_j = 0),\ j = 1, 2, \ldots, n.$ (1.3)

If $\underline{X} = (X_1, X_2, \ldots, X_n)$ is the observed sample on X, then $((X_1, Z_1), (X_2, Z_2), \ldots, (X_n, Z_n))$ becomes the complete sample as $(X_1, X_2, \ldots, X_n)$ is augmented by $(Z_1, Z_2, \ldots, Z_n)$. If $X_j > s$, then $Z_j = 1$ and if $X_j \le s$, then $Z_j = 0$ or 1. In other words, we have no information on Z_j for $X_j \le s$. Hence, $\{Z_j: X_j \le s\}$ can be treated as the missing data.

The likelihood function corresponding to the complete data is then given by

$$L_c(\theta, \varphi | \underline{x}, \underline{u}) = \prod_{j=1}^{n} \left((1-\varphi) \frac{\theta}{\beta} (\theta x_j)^{\beta-1} \exp(-\theta x_j) \right)^{u_j} \left(\frac{\varphi}{s} \right)^{1-u_j},$$

where $u_j = 1$ if $X_j > s$ and $u_j = Z_j$, if $X_j \le s$. Consequently, the log-likelihood can be written as

$$\log L_c(\theta,\varphi|\underline{x},\underline{u}) = \sum_{j:x_j>s}\left[\log(1-\varphi)+\log\theta-\log\beta+(\beta-1)\left(\log\theta+\log x_j\right)-\theta x_j\right]$$
$$+\sum_{j:x_j\le s}u_j\left[\log(1-\varphi)+\log\theta-\log\beta+(\beta-1)\left(\log\theta+\log x_j\right)-\theta x_j\right]$$
$$+\sum_{j:x_j\le s}(1-u_j)[\log\varphi-\log s]. \quad (1.4)$$

The E – Step: By taking the expectation of the log-likelihood, we get

$$E[\log L_c(\theta,\varphi|\underline{x},\underline{u})] = \sum_{j:x_j>s}\left[\log(1-\varphi)+\log\theta-\log\beta+(\beta-1)\left(\log\theta+\log x_j\right)-\theta x_j\right]$$
$$+\sum_{j:x_j\le s}E(Z_j)\left[\log(1-\varphi)+\log\theta-\log\beta+(\beta-1)\left(\log\theta+\log x_j\right)-\theta x_j\right].$$
$$+\sum_{j:x_j\le s}(1-E(Z_j))[\log\varphi-\log s]$$

Now, $E(Z_j)$ is replaced by the conditional expectation $E(Z_j|\theta_0,\varphi_0,x_j)$, where θ_0 and φ_0 are initial estimates of θ and φ.

$$E(Z_j|\theta_0,\varphi_0,x_j) = 0.P(Z_j=0|\theta_0,\varphi_0,x_j)+1.P(Z_j=1|\theta_0,\varphi_0,x_j)$$

$$= P(Z_j=1|\theta_0,\varphi_0,x_j).$$

Using the Bayes' Theorem,

$$P(Z_j=1|\theta_0,\varphi_0,x_j) = \frac{P(A_j|Z_j=1)P(Z_j=1)}{P(A_j|Z_j=0)P(Z_j=0)+P(A_j|Z_j=1)P(Z_j=1)},$$

where $A_j = [X_j = x_j \mid \theta_0,\varphi_0]$.

That is

$$P(Z_j=1|\theta_0,\varphi_0,x_j) = \frac{(1-\varphi_0)\dfrac{\theta_0}{\beta}(\theta_0 x_j)^{\beta-1}\exp(-\theta_0 x_j)}{\dfrac{\varphi_0}{s}+(1-\varphi_0)\dfrac{\theta_0}{\beta}(\theta_0 x_j)^{\beta-1}\exp(-\theta_0 x_j)} = w_j, say,\ for\ j: x_j \le s.$$

Therefore,

$$E[\log L_c(\theta,\varphi|\underline{x},\underline{u})] = \sum_{j:x_j>s}\left[\log(1-\varphi)+\log\theta-\log\beta+(\beta-1)\left(\log\theta+\log x_j\right)-\theta x_j\right]$$
$$+\sum_{j:x_j\le s} w_j\left[\log(1-\varphi)+\log\theta-\log\beta+(\beta-1)\left(\log\theta+\log x_j\right)-\theta x_j\right].$$
$$+\sum_{j:x_j\le s}(1-w_j)[\log\varphi-\log s]$$

The M- Step: In this step, $E[\log(L_c(\theta,\varphi\,|\,\theta_0,\varphi_0,\underline{x},\underline{u})]$ is maximized for θ and φ. If θ_1 and φ_1 are the values of θ and φ which maximize $E[\log(L_c(\theta,\varphi\,|\,\theta_0,\varphi_0,\underline{x},\underline{u})]$, then the E – step is repeated using θ_1 and φ_1. After each of the iterations, the value of the likelihood $L(\theta,\varphi\,|\,\underline{x})$ specified in (1.4) can be evaluated and observed whether it is increasing. The iterative procedure is terminated when $L(\theta,\varphi;\underline{x})$ or $\log L(\theta,\varphi;\underline{x})$ converges to a value correct to a desirable number of decimal places.

Since $E[\log(L_c(\theta,\varphi\,|\,\theta_0,\varphi_0,\underline{x},\underline{u})]$ is differentiable with respect to both θ and φ, the values of θ and φ for which $E[\log(L_c(\theta,\varphi\,|\,\theta_0,\varphi_0,\underline{x},\underline{u})]$ is a maximum can be obtained by the method of calculus.

we get

$$\frac{\partial E[\log(L_c(\theta,\varphi\,|\,\theta_0,\varphi_0,\underline{x},\underline{u})]}{\partial\theta}=0\Rightarrow\sum_{j:x_j>s}\left[\frac{1}{\theta}-x_j+\frac{\beta-1}{\theta}\right]+\sum_{j:x_j\le s}w_j\left[\frac{1}{\theta}-x_j+\frac{\beta-1}{\theta}\right]=0$$

and

$$\frac{\partial E[\log(L_c(\theta,\varphi\,|\,\theta_0,\varphi_0,\underline{x},\underline{u})]}{\partial\varphi}=0\Rightarrow\sum_{j:x_j>s}\left(\frac{-1}{1-\varphi}\right)+\sum_{j:x_j\le s}w_j\left(\frac{-1}{1-\varphi}\right)+\sum_{j:x_j\le s}(1-w_j)\left(\frac{1}{\varphi}\right)=0.$$

By solving

$\dfrac{\partial E[\log(L_c(\theta,\varphi\,|\,\theta_0,\varphi_0,\underline{x},\underline{u})]}{\partial\theta}=0$ *and* $\dfrac{\partial E[\log(L_c(\theta,\varphi\,|\,\theta_0,\varphi_0,\underline{x},\underline{u})]}{\partial\varphi}=0$, we get θ_1 *and* φ_1 by solving the above equations. The n_g and n_l are defined respectively as number of observations that are greater than 's' and less than or equal to 's' respectively.

The EM algorithm for computing the MLEs of the parameters in the perturbed gamma model can be summarized as follows:

(A) Choose the initial estimates θ_0 and φ_0.

(B) Using the realization (x_1, x_2, . . . , x_n) of the observed sample, compute

$$w_j = \frac{(1-\varphi_0)\dfrac{\theta_0}{\beta}(\theta_0\, x_j)^{\beta-1}\exp(-\theta_0 x_j)}{\dfrac{\varphi_0}{s} + (1-\varphi_0)\dfrac{\theta_0}{\beta}(\theta_0\, x_j)^{\beta-1}\exp(-\theta_0 x_j)}, \quad \text{for } j: x_j \le s.$$

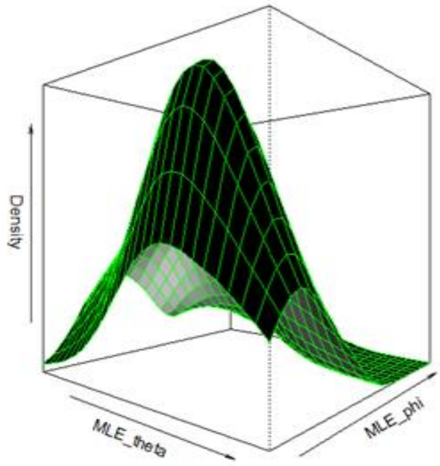
(C) Take $\theta_1 = \dfrac{\beta\left(n_g + \sum_{j:x_j \le s} w_j\right)}{\sum_{j:x_j > s} x_j + \sum_{j:x_j \le s} w_j x_j}$ and $\varphi_1 = \dfrac{n_l - \sum_{j:x_j \le s} w_j}{n}$.

(D) Repeat step (B) by fixing $\theta_0 = \theta_1$ and $\varphi_0 = \varphi_1$ until $L(\theta,\varphi;\underline{x})$ or $\log L(\theta,\varphi;\underline{x})$ converges to a value correct to a desirable number of decimal places.

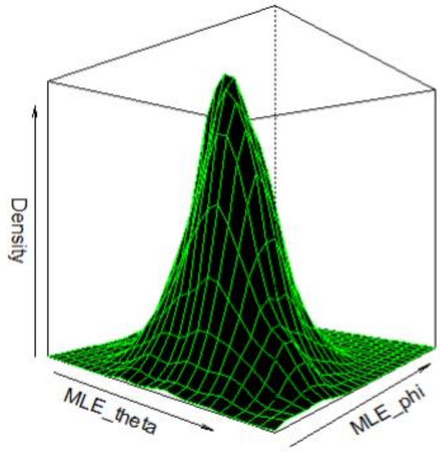
A reasonable initial estimate of □ is $\dfrac{n_l}{n}$ and $\dfrac{n}{\sum_{j=1}^{n} x_j}$ can be taken as an initial estimate of □.

The following figure represents the three-dimensional graph for MLEs of the parameters to demonstrate the asymptotic normality. As the sample size increases the curve approaches to normality.

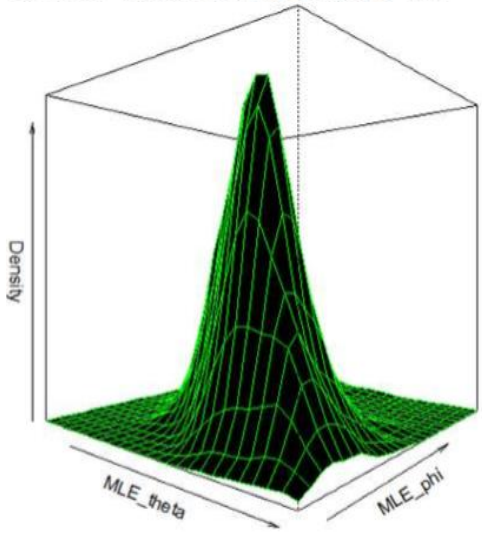
Fig4.1: Bivariate density of MLEs of θ and ϕ for perturbed Gamma model



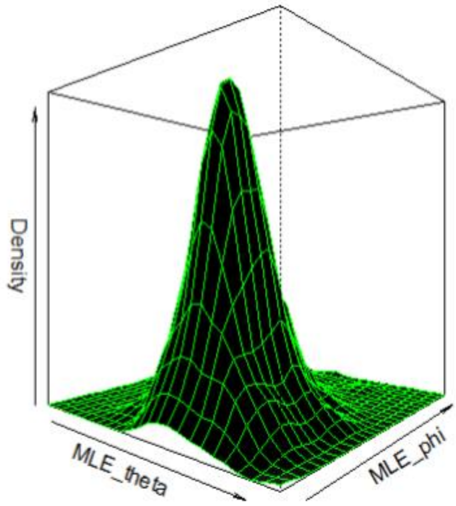
N=10, s=1, beta=1.5, th=0.75, phi=0.4



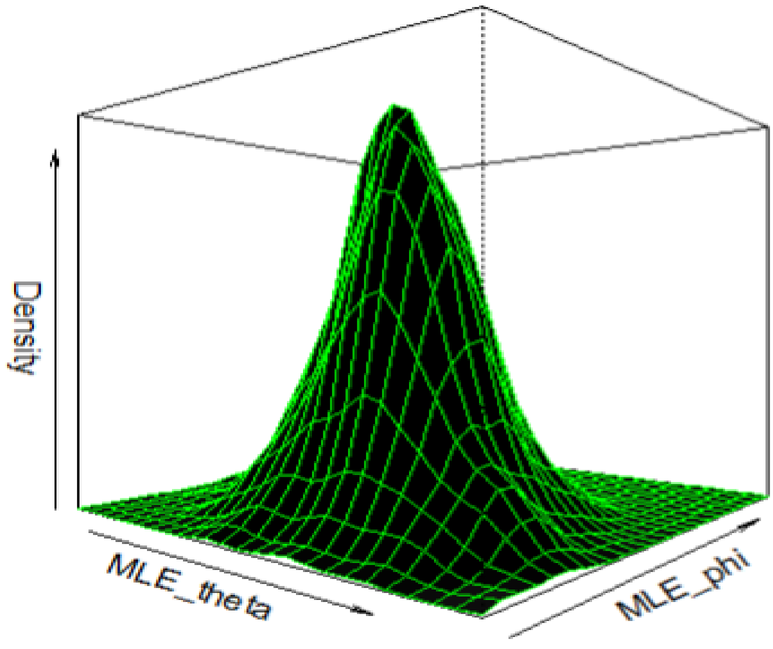
Sample size =200, $\beta = 2, \theta = 0.5, \varphi = 0.2, and\ s = 1$



Sample size =50, $\beta = 1.5, \theta = 0.75, \varphi = 0.4, and\ s = 1$



Sample size =100, $\beta = 2, \theta = 0.5, \varphi = 0.2, and\ s = 1$



Sample size =200, $\beta = 2, \theta = 0.5, \varphi = 0.2, and\ s = 1$

5.1 Conclusion

When the data is missing, the estimation of parameters in the above model using maximum likelihood estimation is very complicated structure and likelihood equation do not yield closed form expression. Even numerical methods like Newton Rapson also tedious, then EM algorithm is used to estimate the parameters of perturbed gamma model. The simulation study has been carried out to demonstrate the bivariate density of MLEs of θ and φ are asymptotic normality of the estimators.

Acknowledgment

I would like to thank my PhD guide and mentor late Prof. Nanjundan G for providing me a constructive comments and suggestions to complete the paper.

References:

1. David A. Van Dyk, Xiao-Li Meng and Donald B. Rubin, (1995). Maximum Likelihood Estimation via the Ecm Algorithm: Computing The Asymptotic Variance, Statistica Sinica 5, 55-75
2. Dempster A.P., N.M. Laird, and D.B. Rubin (1977). Maximum likelihood Estimation from incomplete data via the EM algorithm (with discussion), J. Roy. Statist. Soc. Ser. B, 39,1 -38.
3. Laurent Bordes,Christian Paroissin &Ali Salami (2016). Parametric inference in a perturbed gamma degradation process, HAL Publisher.
4. Massimiliano Giorgio , Agostino Mele , Gianpaolo Pulcini (2018). A perturbed gamma degradation process with degradation dependent non-Gaussian measurement errors,
5. McLachlan G.J. and Krishnan T., (2008). The EM Algorithm and Extensions,2/e, John Wiley, New York.
6. McLachlan, G.J. and Krishnan, T. (1997). The EM Algorithm and Extensions. John Wiley and Sons, New York.
7. McLeish, D.L. and Small, C.G. (1988). Theory and Applications of Statistical Inference Functions", Springer Lecture Notes in Statistics 44.
8. Nanjundan, G and Raveendra Naika, T. (2011) "Maximum Likelihood Estimation of Parameters in a Perturbed Exponential Model", International Journal of Agricultural and Statistical Sciences, Vol.7, No.2, pp. 527-534.
9. Nanjundan, G. (2006). An EM algorithmic approach to maximum likelihood estimation in a mixture model. Vignana Bharathi, Vol. 18, 7 – 13.
10. Nanjundan G and T. Raveendra Naika (2011). Maximum Likelihood Estimation of Parameters in a Perturbed Exponential Model in International Journal of Agricultural and Statistical Sciences, Vol.7, No.2, pp. 527-534 , December 2011 issue.
11. Raveendra Naika T and Sadiq Pasha (2021). Asymptotic normality of the estimators of the parameters in Perturbed Weibull mode using EM algorithm in Advances and Applications in Mathematical Sciences, Volume 20, Issue 10, August 2021, Pages 2261-2268, Mili Publications.