

Modified Moment Estimation for a Two Parameter Gamma Distribution

Emily Kirimi¹, Abel Ouko², Cheruiyot W. Kipkoech³

¹Technical University of Kenya

Department of Mathematics and Statistics

²The East African University, Kenya

Department of Mathematics

³Maasai Mara University, Kenya

Department of Mathematics and Physical Sciences

Abstract: A study of a modification of moment estimation of a two parameter gamma distribution has been carried out. The scale parameter has been estimated using the Modified Moment Estimates (MMEs) and their performance is compared to the Maximum Likelihood Estimates (MLEs). Modified Moment Estimates are easy to compute relative to Maximum Likelihood Estimates. To adjust their bias, the Jackknife technique has been used. The efficiency of these estimators is compared by calculating their respective variances. The R has been used to carry out the Monte Carlo simulation of data. The results obtained show that the Modified Moment Estimates perform as well as the Maximum Likelihood Estimates when bias as well as variation of data is taken into account.

Keywords: Gamma distribution, Modified Moment Estimates, Maximum Likelihood Estimates

I. Introduction

Several authors have considered the problem of estimating the parameter of the gamma distribution. Fisher (1922) showed that the Method of Moments may be inefficient for estimating a two-parameter gamma distribution and suggested the use of Maximum Likelihood (ML) method. Kendall and Stuart (1977) showed that efficiency of the estimated shape parameter (λ) of a gamma distribution by the method of moments may be as low as 22 percent. The main task is to find the modified moment estimators which are close to efficiency and easy to calculate and investigate their performance by comparing them with MLEs. We shall adjust the bias of these estimators to get MLEs and MMEs by use of Jackknife technique.

The gamma distribution is a two parameter family of continuous probability distributions. It has a scale parameter α and a shape parameter λ . The gamma distribution is frequently used as a probability model for waiting times. For instance, in life testing, the waiting time until failure is a random variable that is frequently modeled using the gamma distribution. The gamma distribution can also be used to model components that have two causes of failure such as sudden catastrophic failures and wear out failures.

A random variable X is said to have a gamma distribution with parameters α and λ , if its probability density function is of the form

$$f(x, \alpha, \lambda) = \frac{1}{\alpha \Gamma(\lambda)} \left(\frac{x}{\alpha} \right)^{\lambda-1} e^{-x/\alpha} \quad x > 0$$

Where $\alpha > 0$ is the scale parameter and $\lambda > 0$ is the shape parameter.

The gamma function

$$\Gamma(\lambda) = \int_0^{\infty} x^{\lambda-1} e^{-x} dx$$

is equal to the factorial function $(\lambda - 1)!$, where λ a positive integer. $\Gamma(\lambda)$ can be obtained from the recurrence relation (Hogg and Craig, 1978)

$$\Gamma(\lambda) = (\lambda - 1)\Gamma(\lambda - 1) = (\lambda - 1)!$$

Two-parameter gamma distribution is a popular distribution for analyzing lifetime data. It has lots of applications in different fields other than lifetime distributions. The two parameters of a gamma distribution (scale and shape) and because of these parameters, the distribution has a bit of flexibility to analyze any positive real data. It has increasing as well as decreasing failure rate depending on the shape parameter. There are

specifically three basic shapes of the distribution depending on whether $\lambda < 1$, $\lambda = 1$ and $\lambda > 1$. Gamma hazard function is capable of representing these three shapes.

The sum of independent and identically distributed gamma random variables has a gamma distribution. If a system has one component and n-spare parts and if the component and each spare part have independent and identically distributed gamma lifetime distributions, then the lifetime distribution of the system also follows a gamma distribution.

Table 1. Properties of gamma distribution

Probability Density Function (pdf)	$f(x, \alpha, \lambda) = \frac{1}{\alpha \Gamma(\lambda)} \left(\frac{x}{\alpha}\right)^{\lambda-1} e^{-x/\alpha}, \quad x > 0$
Cumulative Density Function (cdf)	$F(x) = \frac{\Gamma_x(\lambda)}{\Gamma(\lambda)}$ where $\Gamma_x(\lambda) = \int_0^x t^{\lambda-1} e^{-t} dt$
Parameter Restriction	$\alpha > 0 \quad \lambda > 0$
Mean	$\lambda \alpha$
Variance	$\lambda \alpha^2$

II. Jackknife Technique

The basic idea behind Jackknife technique is to systematically re-compute the statistic estimate leaving out one observation at a time. A sample of size n is selected and then the MMEs and MLEs of n-1 observations are computed. Let $\hat{\theta}_j$ be the estimate obtained by leaving out the i^{th} observation and then using the n-1 observations, the Jackknife estimate is given by

$$\hat{\theta}_j = n\hat{\theta} - (n-1)\hat{\theta}_{/i}$$

where

$$\hat{\theta} = \frac{1}{n} \sum_{i=1}^n \hat{\theta}_{/i} \quad \hat{\theta}_{/i} = \frac{1}{n-1} \sum_{j \neq i}^n X_j$$

III. Bias of an estimator

Bias of an estimator is the difference between the expected value of the estimator and the true value of the parameter being estimated. An estimator with zero bias is called unbiased. Otherwise the estimator is said to be biased. If $\hat{\theta}$ is an estimator of the parameter θ , then the bias of an estimator is defined as

$$\text{Bias}(\hat{\theta}) = E(\hat{\theta}) - \theta = E(\hat{\theta} - \theta)$$

Where $E(\hat{\theta})$ is the expected value and θ is the true value

IV. Efficient estimator

Let T1 and T2 be any two unbiased estimators of a parameter $\theta \in \Omega$. We define efficiency of T1 relative to T2

$$\text{by eff}(T1/T2) = \frac{\text{var}(T_2)}{\text{var}(T_1)} \text{ and say that T1 is more efficient than T2 if } \text{eff}(T1/T2) > 1.$$

If T is a most efficient estimator of θ then the efficiency of any other unbiased estimator T1 of θ relative to T is

$$\text{given by } \text{eff}(T1/T) = \frac{\text{var}(T)}{\text{var}(T_1)} \leq 1$$

4.1 Asymptotic efficiency

An unbiased estimator T1 is asymptotically most efficient if $\lim_{n \rightarrow \infty} \text{eff}(T1/T) = 1$ where T is the most efficient estimator of θ .

V. Consistency

A consistent sequence of estimators is a sequence of estimators that converge in probability to the quantity being estimated as the index (usually the size of the sample) grows without bound. In other words, increasing the sample size increases the probability of the estimator being close to the population parameter θ . A sequence of estimators $\{T_n ; n \geq 0\}$ is a consistent estimator for θ if and only if, for every $\epsilon > 0$, no matter how small, we have

$$\lim_{n \rightarrow \infty} Pr \left\{ |T_n - \theta| < \epsilon \right\} = 1$$

VI. Monte Carlo Simulation

Monte Carlo Simulation is a scheme employing numbers for solving problems where passage of time plays no substantive role. Here parameter estimates will be calculated, analyzed and compared with the true parameters of the two parameter gamma distribution.

VII. Maximum likelihood estimation method.

The two parameter gamma distribution has a density function

$$f(x, \alpha, \lambda) = \frac{1}{\alpha \Gamma(\lambda)} \left(\frac{x}{\alpha}\right)^{\lambda-1} e^{-x/\alpha} \quad x > 0 \quad \lambda, \alpha > 0. \dots\dots\dots (i)$$

Let x_1, x_2, \dots, x_n ($n > 1$) represent a random sample of values of X

The log likelihood function is given by

$$l = n \left\{ -\lambda \log \alpha - \log \Gamma(\lambda) \right\} + (\lambda - 1) \sum_{i=1}^n \log x_i - \frac{1}{\alpha} \sum_{i=1}^n x_i \dots\dots\dots (ii)$$

Taking partial derivatives with respect to α and λ

$$\frac{\partial l}{\partial \lambda} = M(\alpha, \lambda) = n \left\{ -\log \alpha - \Psi(\lambda) \right\} + \sum_{i=1}^n \log x_i \dots\dots\dots (iii)$$

$$\frac{\partial l}{\partial \alpha} = g(\alpha, \lambda) = n \left\{ \frac{-\lambda}{\alpha} \right\} + \frac{1}{\alpha^2} \sum_{i=1}^n x_i \dots\dots\dots (iv)$$

Where $\Psi(\lambda) = \frac{d}{d\lambda} [\log \Gamma(\lambda)]$ is the digamma function

The maximum likelihood estimates (MLEs) of the parameters of the gamma distribution are the values of the parameter that produce the maximum joint probability density for the observed data X. The gamma distribution is a regular unimodal distribution and its MLEs are derived by finding the locations of the function peaks with the estimated parameters (α, λ) .

Equations (iii) and (iv) gives the MLEs as $\hat{\alpha} = \frac{\bar{x}}{\hat{\lambda}}$ and $Y = \ln \hat{\lambda} - \Psi(\hat{\lambda})$

where $Y = \overline{(\ln x - \ln \bar{x})}$

Analytically the explicit expression for the MLEs cannot be obtained and their solutions require numerical techniques. Two methods of numerical solutions have been developed.

7.1 Newton-Raphson Method

Solving equation (iv) for α we get

$$\alpha = \frac{\bar{x}}{\lambda} \dots\dots\dots (v)$$

\bar{x} denoting the sample mean.

Substituting this into the second equation to get an expression in terms of λ only we get

$$-n \ln \left(\frac{\bar{x}}{\lambda} \right) - n \Psi(\lambda) + \sum_{i=1}^n \ln(x_i) = 0$$

This can be rearranged to give

$$\ln(\lambda) - \Psi(\lambda) - \ln \left(\frac{\bar{x}}{x} \right) = 0 \dots \dots \dots (vi)$$

$\tilde{x} = (\prod_{i=1}^n x_i)^{1/n}$ denoting the geometric mean

Equation (vi) can be solved numerically using Newton Raphson iterations. The MLEs are therefore given by equation (v) and the solution of equation (vi). This method has been explored by Choi and Wette (1969)

7.2 MLE Scoring Method.

Choi and Wette (1969) considered a maximum likelihood scoring method and noted that although this proposed technique involves more computation than the Newton Raphson method, it is appealing because it provides a statistical criterion for stopping the iterations.

VIII. Method of Moment estimators.

The method of moments consists of equating sample moments with unobservable population quantities for the quantities to be estimated. Suppose X_1, X_2, \dots, X_n are independent identically distributed random variables with a gamma pdf

$$f(x, \alpha, \lambda) = \frac{1}{\alpha \Gamma(\lambda)} \left(\frac{x}{\alpha} \right)^{\lambda-1} e^{-x/\alpha} \quad \forall x > 0$$

Considering the moment generating function for the gamma distribution we get the following.

$$\begin{aligned} MX(t) &= E(e^{tx}) \\ &= \int_0^{\infty} e^{tx} \frac{1}{\alpha^\lambda \Gamma(\lambda)} x^{\lambda-1} \exp\left(-\frac{x}{\alpha}\right) dx = \alpha^{-\lambda} \left(\frac{1}{\alpha} - t\right)^{-\lambda} \end{aligned}$$

Computing the population moments for the gamma we have

$$\begin{aligned} U_1(\alpha, \lambda) &= E[X] \\ &= \alpha \lambda \end{aligned}$$

And the second population moment is

$$\begin{aligned} U_2(\alpha, \lambda) &= E[X^2] \\ &= \lambda \alpha^2 (\lambda + 1) \end{aligned}$$

These are the first two population moments. Let the first and second sample moments are M_1 and M_2 respectively. Equating the population moments with the sample moments we get

$$\begin{aligned} M_1 &= \alpha \lambda \\ M_2 &= \alpha^2 \lambda (\lambda + 1) \end{aligned}$$

where $M_1 = \frac{1}{n} \sum_{i=1}^n x_i$

$$M_2 = \frac{1}{n} \sum_{i=1}^n x_i^2$$

Solving for α and λ we get the following as the method of moment estimators of α and λ

$$\hat{\lambda} = \frac{M_1^2}{M_2 - M_1^2}$$

where

$$f1(y) = (.5000876 + .1648852y - .0544274y^2)/y$$

$$f2(y) = (8.898919 + 9.059950y - .9775373y^2)/(17.79728 + 11.968477y + y^2) y$$

To obtain $\hat{\lambda}_{ML}$ often starts with the standard method of moment estimates

$$\hat{\lambda}_{ML} = \bar{X}^2 / s^2 \dots\dots\dots(x)$$

$$\hat{\alpha} = s^2 / \bar{X} \dots\dots\dots(xi)$$

where s^2 is the sample variance

These estimates arise from the following observation. From the Method of Moments it follows that

$$\hat{\alpha} = \frac{M_2 - M_1^2}{M_1}$$

$$\hat{\lambda} = \frac{M_1^2}{M_2 - M_1^2}$$

From MLEs we got the following

$$\hat{\alpha} = \frac{\bar{X}}{\lambda} = \frac{E(X)}{\lambda}$$

$$\hat{\lambda} = \frac{E(X)^2}{E(X^2) - [E(X)]^2} = \frac{E(X^2)}{VarX}$$

If $X \rightarrow G(\alpha, \lambda)$ then

$$\hat{\lambda} = \frac{E[X]^2}{Var[X]} \quad \text{and} \quad \hat{\alpha} = \frac{E[X]}{\lambda} \dots\dots\dots(xii)$$

Generally for $K > 0$ we get $\lambda = \frac{E[X^k]E[X]}{Cor\left[\frac{X^K}{K}, X\right]} \dots\dots\dots(xiii)$

If $K=1$ then we get equation (xii) which employs two moments, otherwise equation (xiii) employs three moments as noted by Wiens et al (2003) who also did a passage to the limit as k tend to zero and obtained

$$\hat{\lambda} = \frac{E[X]}{Cov[\ln X, X]} \dots\dots\dots(xiv)$$

Generally for $K > 1$ we employ higher moments and get the estimates as follows

$$\hat{\alpha}_k = E[X] / \hat{\lambda}_k \quad \hat{\lambda}_k = \frac{E[X^k]E[X]}{Cor\left[\frac{X^K}{K}, X\right]} \quad \text{for } K > 0$$

For $K=1$ we propose the modified moment estimators as follows

$$\hat{\alpha}_1 = \frac{E[X]}{\hat{\lambda}_1} \quad \hat{\lambda}_1 = \frac{E(X)^2}{var X}$$

X. Simulation

Simulation experiment was carried out to compare the performance of the Modified Moment Estimators and Maximum Likelihood Estimators for the two-parameter gamma distribution. The Monte Carlo simulation was performed for the samples of sizes $n=10, 30, 50, 100$ and generated values of α and λ . To reduce the bias the Jackknife technique was applied to the simulated values.

XI. Results

Table 2: Constrained estimates of $\alpha=1.0$ when $\lambda=1, 3$ and 6

λ	n	MME	MLE	JMME	JMLE	Var-MME	Var-MLE
1	10	1.3784506	0.4171561	1.015763	0.918888	7.05E-04	0.185045
	30	0.9730263	0.6632311	1.025191	0.913166	0.002344	0.024809
	50	1.0784946	1.3314015	1.022739	0.897627	1.84E-03	0.003893
	100	1.0784946	0.9853911	1.021939	0.905674	4.04E-03	0.008338
3	10	0.795372	1.031961	0.904591	0.904591	5.33E-03	0.003236
	30	1.041266	0.709764	1.023604	0.912084	1.30E-03	0.004767
	50	0.9797	1.045034	1.025036	0.904287	1.76E-04	0.016748
	100	0.982463	0.862275	1.024972	0.908537	3.39E-03	0.006422
6	10	0.883375	0.356838	1.027276	0.920291	-	-
	30	0.939706	0.82745	1.025966	0.909347	-	-
	50	1.087375	1.003111	1.022532	0.905262	-	-
	100	0.977695	1.099507	1.025083	0.90302	-	-

- MME** - Modified Moment Estimates
- MLE** - Maximum Likelihood Estimates
- JMME** - Jackknife Modified Moment Estimates
- JMLE** - Jackknife Maximum Likelihood Estimates
- Var-MME** - Variance of Modified Moment Estimates
- Var-MLE** - Variance of Maximum Likelihood Estimates

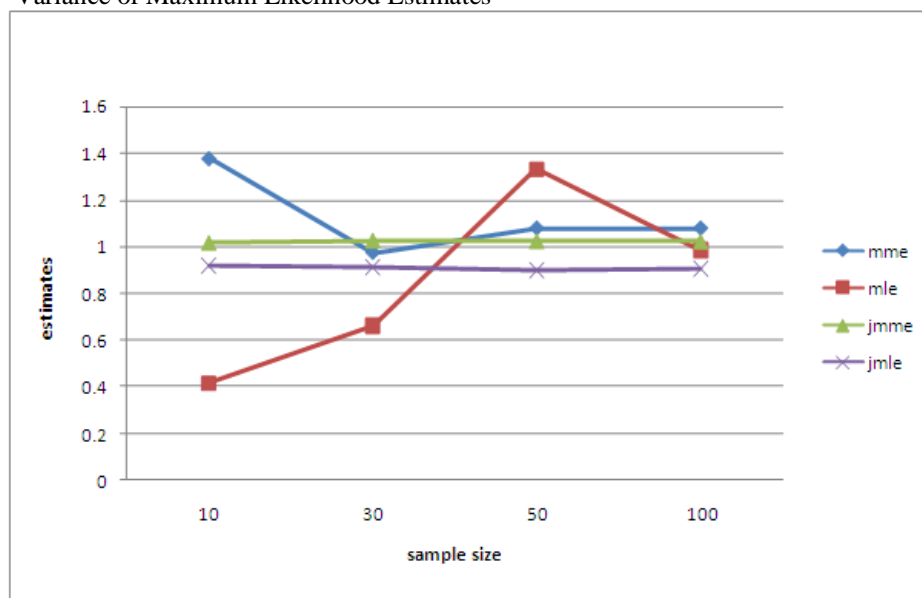


Figure 1. Estimates of $\alpha=1$ for $\lambda=1$

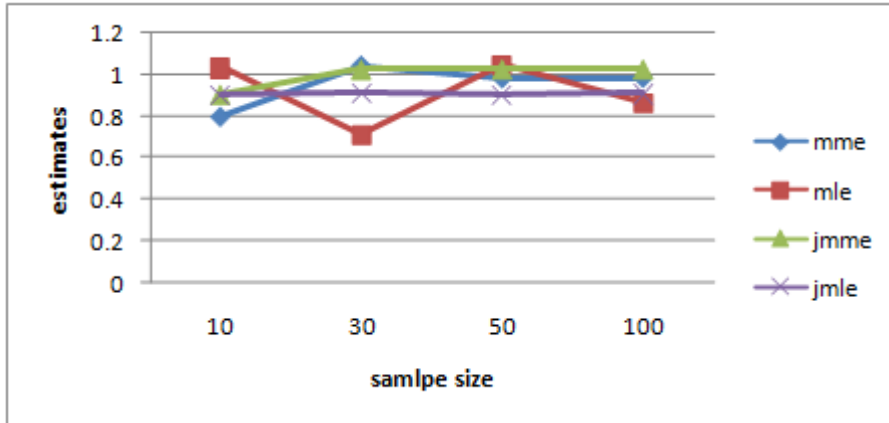


Figure 2. Estimates of $\alpha=1$ for $\lambda=3$

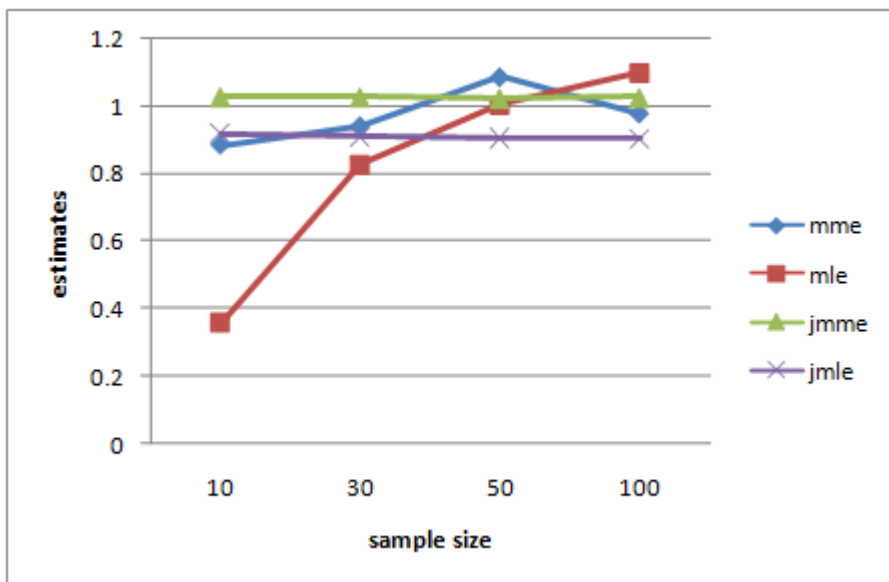


Figure 3. Estimates of $\alpha=1$ for $\lambda=6$

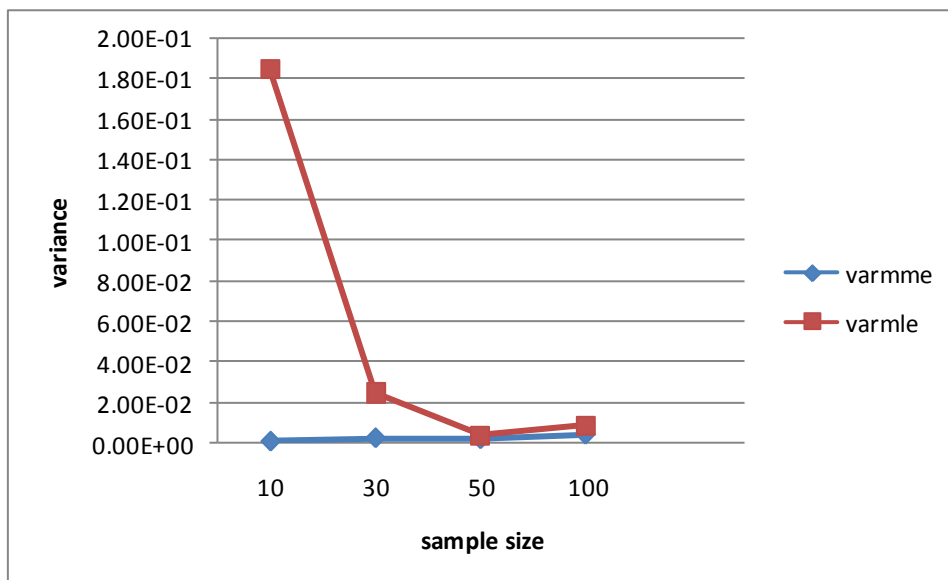


Figure 4. Variance of estimates of $\alpha=1$ for $\lambda=1$

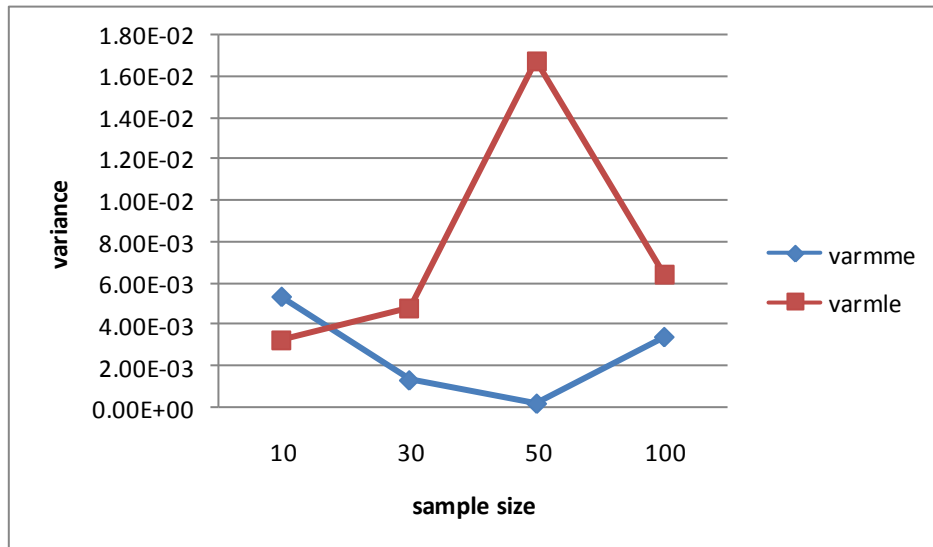


Figure 5. Variance for estimates of $\alpha=1$ for $\lambda=3$

Table 2 shows the estimates of α based on different sample sizes. It is evident that when the sample size is small (say 10) simulated at, the estimates of $\alpha=1$ gives values for both MMEs and MLEs. The estimates of MMEs and MLEs are so close and coincide with small negligible differences as the sample size increases. It also shows how the estimates of α perform as $\lambda=3$ for various sample sizes. The MMEs asymptotically converge to the true value. It is evident that the estimates of α based on MLEs are more biased as compared to the MMEs. The Jackknife bias reduction method works very well in this case. From the simulated results, it is clear that the performance of MMEs are almost for different sizes is desirable. The variances of MMEs are so small as compared to the MLEs. It is evident that the MMEs are highly favorable as estimates of a two parameter gamma distribution. The table also shows the estimates of α perform as $\lambda=6$ for various sample sizes. For large sample sizes the MMEs almost converge to the true value. It also shows the variance of the estimates of α for $\lambda=1$ and $\lambda=3$. It is clear that the variance of the MMEs is small as compared to the MLE counterpart.

XII. Conclusion

The classical approach to estimating the parameters of the two parameter gamma distribution has largely centered on approximating the solution in the likelihood equation. Here the estimates of the two parameter gamma distribution have been obtained using modified moment estimation, which requires no numerical iterations and are easy to implement. The biases of MMEs and MLEs were adjusted using Jackknife technique. The estimates of MMEs work well as the MLEs especially when bias is taken into account for different sample sizes.

References

- [1]. Choi, S.C and Wette R., 1969. Maximum likelihood estimation of parameters of the Gamma Distribution and their bias. *Technometrics* 11, 83-90
- [2]. Fisher, R. A., 1922. On the mathematical foundations of theoretical statistics, *Philosophical transactions of the Royal Society of London A222*, 309-368.
- [3]. Greenwood, J.A and Durand, D., 1960. Aids for fitting the gamma distribution by maximum likelihood. *Technometrics*, 2, 55-65
- [4]. Hogg, R.V and Craig, A.T., 1978. *Introduction to mathematical statistics*. 4th edition, New York, Macmillan.
- [5]. Kendall, M and Stuart, A., 1977. *The advanced theory of statistics*. Charles Griffin and company limited.
- [6]. Thorn, H.C., 1958. A note on the gamma distribution, *Monthly Weather Review*, 86, 117-122
- [7]. Wiens, D.P. Cheng and Beaulieu N.C (2003), A class of methods of moments estimators for the two-parameter gamma family. *Pak.J.Statistics* 19, 129-141.